The use of 18FDG PET in NSCLC

Herder, G.J.M.

2006

document version
Publisher's PDF, also known as Version of record

The use of 18FDG PET in NSCLC

The use of 18FDG PET in NSCLC

Part of this thesis was supported by ZonMw Program for Health Care Efficiency Research

The publication of this thesis was financially supported by
Amersham Health B.V., Amgen B.V. Breda, AstraZeneca B.V., B.V. Cyclotron VU, Eli Lilly, GE Healthcare, GlaxoSMithKline, Janssen-Cilag B.V., Novartis Pharma B.V., Ortho Biotech a division of Janssen-Cilag B.V., Paes Nederland-Olympus, Roche, sanofi-aventis, Tyco Healthcare/Mallinckrodt, Veenstra Instruments B.V.

Cover design:	Alex de Jong
Lay-out:	Chris Bor, Medische Fotografie en Illustratie, AMC
Printing:	Buijten & Schipperheijn, Amsterdam, The Netherlands
ISBN:	90-810478-1-7

VRIJE UNIVERSITEIT

The use of 18FDG PET in NSCLC

ACADEMISCH PROEFSCHRIFT

ter verkrijging van de graad Doctor aan
de Vrije Universiteit Amsterdam,
op gezag van de rector magnificus
prof.dr. T. Sminia,
in het openbaar te verdedigen
ten overstaan van de promotiecommissie
van de faculteit der Geneeskunde
op woensdag 26 april 2006 om 13.45 uur
in de aula van de universiteit,
De Boelelaan 1105

door
Gerarda Johanna Maria Herder

geboren te Amersfoort

promotoren:	prof.dr. P.E. Postmus
	prof.dr. G.J.J. Teule
copromotoren:	prof.dr. O.S. Hoekstra
	prof.dr. E.F. Smit

Voor mijn ouders

Contents

Chapter 1	Introduction	9
Chapter 2	Practice, efficacy and cost of staging suspected non-small cell lung cancer: a retrospective study in two dutch hospitals	17
Chapter 3	The performance of ^{18}F-fluorodeoxyglucose positron emission tomography (^{18}FDG PET) in small solitary pulmonary nodules	27
Chapter 4	Clinical prediction model to characterise pulmonary nodules: validation and added value of ^{18}FDG PET	39
Chapter 5	Prospective use of serial questionnaires to evaluate the therapeutic efficacy of ^{18}FDG PET in (suspected) lung cancer	51
Chapter 6	Staging of non-small cell lung cancer and application of ^{18}FDG PET: a cost modeling approach	63
Chapter 7	Traditional versus up-front ^{18}FDG PET staging of non-small cell lung cancer (NSCLC): a dutch co-operative randomised study	77
Summary and discussion		91
Nederlandse samenvatting		103
Met andere woorden... (voor niet ingewijden)		113
Dankwoord		117
Abbreviations		123

Chapter 1

Introduction and outline of the thesis

Epidemiology

For lung cancer, routine mortality statistics have confirmed the clinical impression that the disease became more frequent during the first half of the 20th century. Changes in tobacco smoking resulted in changes in the frequency of lung cancer. Tobacco smoking is well established as the main cause of lung cancer and about 90% of cases are thought to be tobacco related. It is the most commonly diagnosed cancer worldwide with 1.35 million new cases in 2002, representing 12.4% of all new cancers. It was also the most common cause of death from cancer, with 1.18 million deaths, or 17.6% of the world total.[1]
In the Netherlands lung cancer is still one of the most common cancer in males. On the basis of the 2000 figures of the Netherlands, the incidence of lung cancer was 79.8 in men and 27 in women per 100.000 per year. Males who have not died from other causes before the age of 75 run a 6.72 percent risk of developing lung cancer before the age of 75 (vs. 2.47% in females).[2] Malignant lung tumours are grossly categorised into two groups based on their morphologic characteristics: small cell lung cancer (SCLC) and non-small cell lung cancer (NSCLC). The majority of lung cancer patients have tumours histologically classified as NSCLC (84%).[3]

Prognosis of NSCLC

Despite improved knowledge of therapy and treatment strategies, the prognosis of patients with NSCLC has not improved significantly over the last decades. Depending on the results of staging procedures, curative surgery is attempted in 30-50%. Nevertheless, survival rates for so-called 'resectable' lung cancer (stage I-IIIA) vary between 15 and 80% at five years. It is estimated that at presentation about 30% of patients has locoregionally advanced disease and about 40% has disseminated disease. The prognosis of unresectable lung cancer is poor. The 1-year survival rate varies between 10-15%, the 5-year survival of advanced disease approaches zero (Table 1).

Clinical evaluation of lung cancer

Patients suspected of having lung cancer undergo several diagnostic procedures to achieve a histological diagnosis and final stage assignment. The modality selected to diagnose lung cancer is based on size and location of the primary tumour, the presence of metastatic spread, and the anticipated treatment plan. The main goals in selecting a specific diagnostic modality are: (1) to maximise the yield of the selected procedure for both diagnosis and staging; and (2) to avoid unnecessary invasive tests for the patient, with special attention to the projected treatment plan.
Determining the diagnosis of the primary tumour, the extent of spread to regional or distant lymph nodes or to other metastatic sites follows two major routes. The first involves obtaining tissue and the measurement of extent of disease which incorporates uniform anatomic criteria of the international tumour, node, metastasis (TNM) staging system (Table 1).[4] It includes

endoscopical procedures for tissue procurement either from bronchial lesions or suspected metastases, imaging studies for localisation of abnormalities and sometimes invasive procedures to assess resectability of the tumour. The second route involves assessment of

Table 1. Tumour staging adapted from Mountain.[4]

Tumour stage				
Tx	Presence of malignant cells, no tumour visible			
T0	No primary tumour			
TIS	Carcinoma in situ			
T1	Tumour < 3.0 cm in size, no invasion beyond the lobar bronchus			
T2	Tumour > 3.0 cm in size; distal atelectasis; involvement of the main bronchus > 2.0 cm distal to the carina; or invades the visceral pleura			
T3	Tumour of any size that invades the chest wall, diaphragm, mediastinal pleura, pericard, or main bronchus < 2.0 cm distal to the main bronchus; or associated atelectasis of the entire lung			
T4	Tumour of any size that invades medidiastinum, the heart, great vessels, esophagus, vertebral body, or carina; the presence of malignant pleural/pericard effusion; or separate tumour nodules in the same lobe			
Nodal Stage				
NX	Regional lymph nodes cannot be assessed			
N0	No lymph node metastasis			
N1	Ipsilateral peribronchial or hilar lymph nodes			
N2	Involvement of ipsolateral mediastinal and/or subcarinal lymph nodes			
N3	Involvement of contralateral or supraclavicular lymph nodes			
Metastastatic stage				
MX	Distant metastasis cannot be assessed			
M0	No evidence of metastasis			
M1	Distant metastasis present			
TNM stage		Cumulative % Surviving*		
		1 Year	3 Year	5 Year
Occult	TX, N0, M0			
Stage 0	Tis, N0, M0			
Stage Ia	T1, N0, M0	90	71	61
Stage Ib	T2, N0, M0	72	46	38
Stage IIa	T1, N1, M0	84	42	37
Stage IIb	T2, N1, M0	63	35	26
	T3, N0, M0	54	30	21
Stage IIIa	T3, N1, M0	58	12	9
	T1-3, N2, M0	51	20	13
Stage IIIb	T4, any N, M0	34	8	7
	Any T, N3, M0	32	6	3
Stage IV	Any T, any N, M1	17	2	<1

* Non-small cell lung cancer. Cumulative percent of patients surviving according to clinical staging criteria

the expected cardiopulmonary reserve after resection. This analysis includes laboratory tests, lung function tests, perfusion and ventilation tests, and a cardiovascular assessment. Judgement of risk factors determines the operability of the patient. Staging of lung cancer patients not only provides important prognostic information with regard to survival, but also guides the decision-making process with regard to selection of optimal treatment.[5] Patients with stage IA, IB, IIA, IIB en selected cases with IIIA disease may benefit from surgical resection, in conjuncture with systemic treatment. Patients with stage IIIB, and IV almost never meet the criteria for surgery.

A variety of diagnostic tests pertaining to each aspect of the TNM staging system is available to assist the clinician in achieving a definitive diagnosis and stage of lung cancer. At the same time, the impact and cost of staging lung cancer are augmenting due to the further increase of newly diagnosed NSCLC, and by the introduction of new diagnostic modalities. Adequate staging is important since it will prevent unnecessary major diagnostic procedures and operations. In *Chapter 2* we evaluated the clinical practice, yield and costs of preoperative staging in patients with (suspected) non-small cell lung cancer in an academic and general hospital during a time period of two years (1993/1994). The importance of staging should be clear since unnecessary major diagnostic procedures and therapy like surgery should be avoided. Evidence- and consensus based guidelines for staging have been developed, which are adapted to new therapeutic and diagnostic developments from time to time.[6-8] One of the latter developments is the introduction of molecular imaging technology in oncology using ^{18}F-fluorodeoxyglucose (^{18}FDG) positron emission tomography (PET).

^{18}FDG PET, basic technique

PET is a physiologic imaging technique that uses radiopharmaceuticals produced by labelling metabolic markers such as glucose with the positron-emitting radionuclide fluorine-18 (a positron emitting isotope), which forms ^{18}FDG. ^{18}FDG preferentially accumulates within cells with a high rate of glycolysis and an increased cellular uptake of glucose, due to an increased expression of glucose transport proteins.[9,10] The radiopharmaceutical is typically imaged by coincidence detection of the two 511 KeV photons that are produced by annihilation of the emitted positrons. Dedicated or state-of-the-art PET cameras make use of Bismuth Germanium Oxide (BGO) multiple block detectors, which form a ring that surrounds the patient. By scanning several axial fields (usually between 15 and 25 cm) a whole-body image of the patient can be reconstructed.

The introduction of ^{18}FDG PET as diagnostic tool in the investigation of suspected lesions would make it possible, with a single examination, to decide whether a lesion is malignant or benign, and to stage a patient with suspected lung cancer.

^{18}FDG PET and solitary pulmonary nodules

The availability of various diagnostic algorithms [11,12] indicates that a standard strategy for clinical management of solitary pulmonary nodules (SPN) has not been defined.

Approximately one-third of pulmonary nodules are radiologically indeterminate and of these, one-third of the resected pulmonary nodules are benign.[13-17] With the increased interest in the use of low dose spiral computed tomography (CT) for early lung cancer detection, the number of coincidental SPNs will increase. Furthermore, repeated screening and technical improvements may result in the detection of very small lesions (2-10 mm in diameter) for which establishing a definitive diagnosis may become even more difficult. The published evidence on the accuracy of ^{18}FDG PET in characterising SPNs (lesions smaller than 3 or 4 cm in diameter) consists of over 450 published cases with a mean sensitivity and specificity of ^{18}FDG PET for detecting malignancy were 93.9% and 77.8%, respectively.[18] Finally, sofar there is little information about ^{18}FDG PET performance in nodules ≤ 10 mm, whereas this is an important theme since the introduction of CT screening for lung cancer. In *Chapter 3* we report on the diagnostic accuracy of ^{18}FDG PET in radiologically indeterminate solitary pulmonary nodules (SPN) ≤10 mm.

Different management strategies circulate for patients with SPN. Decisions should be made adequate and timely to permit curative resection of malign lesions or to avoid surgery in benign lesions. A cost-effectiveness analysis proposed a diagnostic approach in SPN which strongly relied upon the result of clinical risk assessment.[19] This probability estimation was based on clinical and radiological parameters [20], but the underlying quantitative risk algorithm still needs to be externally validated. Further, whether and how addition of ^{18}FDG PET results improves this prediction rule still needs to be established. These two issues are evaluated in *chapter 4*.

Staging NSCLC

^{18}FDG PET has been shown to provide useful information in staging of NSCLC. Several case series claimed ^{18}FDG PET to be superior over conventional staging in determining hematogeneous and locoregional spread.[21-26] These findings suggest that ^{18}FDG PET may both improve and simplify the staging process. Moreover, the idea of a single whole body ^{18}FDG PET to establish disease stage rather than a lot of procedures and tests is appealing. Most studies performed evaluate diagnostic accuracy or clinical outcome. To determine the proper role of ^{18}FDG PET and to appreciate its incremental benefits, it should not be evaluated in isolation from other tests. In addition, improved test performance does not necessarily translate into meaningful changes in clinical treatment decisions; nor does altered treatment necessarily translate into improved patient outcomes. Thus, to fully assess a new diagnostic technology, the evaluation must not only assess test performance but also the impact of the test on decision making and clinical outcomes. A way to perform such studies is determining diagnostic and therapeutic plans before and after the application of ^{18}FDG PET by means of questionnaires.[27,28] *Chapter 5* describes patients presenting with a clinical problem in NSCLC and the degree to which the result of the research (^{18}FDG PET) affected the diagnostic understanding and management. This was studied by questionnaires gathering information on intended diagnostic plans and treatment without PET, actual therapy of choice

after PET and post hoc clinical assessment to grade the usefulness in diagnostic understanding and choice of therapy.

Modelling

To fully assess a new diagnostic technology, the evaluation must assess test performance, the impact of the test on decision making and clinical outcomes, such as morbidity and mortality as well as the value of substitution and cost-effectiveness of a new diagnostic tool. Several strategies can be reconstructed to evaluate the value of ^{18}FDG PET. An advantage of a modeling approach is the possibility to vary the level of substitution of PET in the diagnostic work-up process and its consequent impact on the costs. By using clinical data the number of assumptions can be limited. To answer the question of the position of ^{18}FDG PET within the diagnostic procedures for NSCLC, several study designs can be considered: ^{18}FDG PET on top of regular staging, ^{18}FDG PET before invasive staging or ^{18}FDG PET in front of the staging process. *Chapter 6* shows a cost modelling approach of ^{18}FDG PET in NSCLC. Consequently, these findings have been used in the design of a randomised trial.

The question if ^{18}FDG PET could reduce the number of unnecessary thoracotomies was addressed in the PLUS trial [29] where the addition of ^{18}FDG PET before surgery (after conventional work-up) resulted in a prevention of unnecessary thoracotomies in one out of five patients. However it still remains to be determined whether the use of ^{18}FDG PET could replace conventional work-up (value of substitution) and shorten the work-up period without losing accuracy. In *Chapter 7* we describe our randomised controlled trial of PET immediately after first presentation of patients with (suspected) NSCLC compared to the traditional strategy in routine clinical setting.

Reference List

1. Parkin DM, Bray F, Ferlay J et al. Global cancer statistics, 2002. CA Cancer J Clin 2005;55:74-108
2. Visser O, Coebergh JWW, Schouten LJ et al. Global cancer statistics. Incidence of Cancer in the Netherlands 1999/2000.10.2004. Netherlands cancer registry.
3. Visser O, Siesling S, and Dijck JAAM. Incidence of cancer in the Netherlands 1999/2000. 11. 200. www.ikcnet.nl, Netherlands Cancer Registry (NCR).
4. Mountain CF. Revisions in the International System for Staging Lung Cancer. Chest 1997;111:1710-1717
5. Miller JD, Gorenstein LA, Patterson GA. Staging: the key to rational management of lung cancer. Annals of Thoracic Surgery 1992;53:170-178
6. Duhaylongsod FG, Lowe VJ, Patz EFJ et al. Lung tumor growth correlates with glucose metabolism measured by fluoride-18 fluorodeoxyglucose positron emission tomography. Annals of Thoracic Surgery 1995;60:1348-1352
7. Goldstraw P, Rocmans P, Ball D et al. Pretreatment minimal staging for non-small cell lung cancer: an updated consensus report. Lung cancer 1994;11:S1-S4
8. van Zandwijk N. [Consensus conference on the diagnosis of lung carcinoma]. Ned Tijdschr Geneeskd 1991;135:1915-1919

9. Nolop KB, Rhodes CG, Brudin LH et al. Glucose utilization in vivo by human pulmonary neoplasms. Cancer 1987;60:2682-2689
10. Yamamoto T, Seino Y, Fukumoto H et al. Over-expression of facilitative glucose transporter genes in human cancer. Biochem Biophys Res Commun 1990;170:223-230
11. Ost D, Fein A. Evaluation and management of the solitary pulmonary nodule. Am J Respir Crit Care Med 2000;162:782-787
12. Swensen SJ, Jett JR, Payne WS et al. An integrated approach to evaluation of the solitary pulmonary nodule. Mayo Clin.Proc. 1990;65:173-186
13. Lillington GA. Pulmonary nodules: solitary and multiple. Clin Chest Med 1982;3:361-367
14. Huston J 3rd, Muhm JR. Solitary pulmonary opacities: plain tomography. Radiology 1987;163:481-485
15. Keagy BA., Starek PJ, Murray GF et al. Major pulmonary resection for suspected but unconfirmed malignancy. Ann Thorac Surg 1984;38:314-316
16. Mack MJ, Hazelrigg SR, Landreneau RJ et al. Thoracoscopy for the diagnosis of the indeterminate solitary pulmonary nodule. Ann Thorac Surg 1993;56:825-830
17. Siegelman SS, Zerhouni EA, Leo FP et al. CT of the solitary pulmonary nodule. Am J Roentgenol 1980;135:1-13
18. Gould MK, Maclean CC, Kuschner WG et al. Accuracy of positron emission tomography for diagnosis of pulmonary nodules and mass lesions: a meta-analysis. JAMA 2001;285:914-924
19. Gould MK, Sanders GD, Barnett PG et al. Cost-effectiveness of alternative management strategies for patients with solitary pulmonary nodules. Ann Intern Med 2003;138:724-735
20. Swensen SJ, Silverstein MD, Ilstrup DM et al. The probability of malignancy in solitary pulmonary nodules. Application to small radiologically indeterminate nodules. Arch Intern Med 1997;157:849-855
21. Boland GW, Goldberg MA, Lee MJ et al. Indeterminate adrenal mass in patients with cancer: evaluation at PET with 2-[F-18]-fluoro-2-deoxy-D-glucose. Chest 1995;194:131-134
22. Bury T, Barreto A, Daenen F et al. Fluorine-18 deoxyglucose positron emission tomography for the detection of bone metastases in patients with non-small cell lung cancer. European Journal of Nuclear Medicine 1998;25:1244-1247
23. Dwamena BA, Sonnad SS, Angobaldo JO et al. Metastases from non-small cell lung cancer: mediastinal staging in the 1990s–meta-analytic comparison of PET and CT. Chest 1999;213:530-536
24. Erasmus JJ, Patz EFJ, McAdams HP et al. Evaluation of adrenal masses in patients with bronchogenic carcinoma using 18F-fluorodeoxyglucose positron emission tomography. Am J Roentgenol 1997;1357-1360
25. Pieterman RM, van Putten JW, Meuzelaar JJ et al. Preoperative staging of non-small-cell lung cancer with positron- emission tomography. N Engl J Med 2000;343:254-261
26. Vansteenkiste JF, Stroobants SG, De Leyn PR et al. Mediastinal lymph node staging with FDG-PET scan in patients with potentially operable non-small cell lung cancer: a prospective analysis of 50 cases. Leuven Lung Cancer Group. Chest 1997;112:1480-1486
27. Wittenberg J, Fineberg HV, Black EB et al. Clinical efficacy of computed body tomography. Am J Roentgenol 1978;131:5-14
28. Wittenberg J, Fineberg HV, Ferrucci JT et al. Clinical efficacy of computed body tomography, II. Am J Roentgenol 1980;134:1111-1120
29. Van Tinteren H, Hoekstra OS, Smit EF et al. Effectiveness of positron emission tomography in the preoperative assessment of patients with suspected non-small-cell lung cancer: the PLUS multicentre randomised trial. Lancet 2002;359:1388-1393

Chapter 2

Practice, efficacy and cost of staging suspected non-small cell lung cancer: A retrospective study in two dutch hospitals

Gerarda JM Herder
Paul Verboom
Egbert F Smit
Piet CM van Velthoven
Jan HAM van den Bergh
Cas D Colder
Johan C van Mourik
Pieter E Postmus
Gerrit JJ Teule
Otto S Hoekstra

Thorax 2002;57:11-14

Abstract

Background: A study was undertaken to investigate the clinical practice, yield, and costs of preoperative staging in patients with suspected NSCLC and to obtain baseline data for prospective studies on the cost effectiveness of ^{18}F-fluorodeoxyglucose positron emission tomography in the management of these patients.

Methods: A retrospective study of the medical records of all patients with suspected NSCLC was performed during a 2 year interval (1993-1994) in an academic and a large community hospital.

Results: Three hundred and ninety five patients with suspected NSCLC were identified; 58 were deemed to be medically inoperable and 337 patients proceeded to the staging process. Staging required a mean (SD) of 5.1 (1.5) diagnostic tests per patient (excluding thoracotomy) carried out over a median period of 20 days (IQR 10-31). Many of the tests (including both invasive and non-invasive) were done because previous imaging tests had suggested metastases, and in most cases the results of initial tests proved to be false positive. After clinical staging, 168 patients were considered to be resectable (stage I/II), and 144 patients underwent surgery with curative intent. At surgery 33 patients (23% of those who underwent surgery) were found to have irresectable lesions and 19 patients (13%) had a benign lesion. Surgery was also considered as futile in 22 patients (15%) who developed metastases or local recurrence within 12 months following radical surgery. Hospital admission was responsible for most of the costs.

Conclusion: In many patients staging involved considerable effort in terms of the number of diagnostic tests, the duration of the staging period and the cost, with limited success in preventing futile surgery. Failures relate to the quality of diagnostic preparation at every level of the TNM staging system.

Introduction

The selection of candidates for appropriate treatment, particularly for curative surgery, is the key issue in staging patients with non-small cell lung cancer (NSCLC).[1] A battery of diagnostic tests pertaining to each aspect of the TNM (tumour node metastases) staging system is potentially available. On the other hand, there are societal concerns and economic restraints. The cost of staging lung cancer is increasing because of an increase in newly diagnosed cases of NSCLC. Not surprisingly, guidelines for staging have been developed which can be adapted to include new therapeutic and diagnostic developments.[2-4] In the past decade new diagnostic methods such as magnetic resonance imaging (MRI) and positron emission tomography (PET) have emerged, claiming a role in the staging process.[5] The initial step in assessment of such new technologies should be the evaluation of prevailing clinical practice and its residual inefficiency. [6] Evaluation of the prevailing clinical practice based simply on the estimated use of guidelines is less desirable since adherence to such guidelines is unpredictable.[7] Studying actual patient data may be more accurate than questionnaires [8] or medical audits [9]; it also avoids the potential selection bias of socially desirable answers and accounts better for heterogeneity.

We have systematically investigated the manner, yield and costs of NSCLC staging as carried out in two major Dutch hospitals. In addition, this study provided baseline data for prospective studies on the cost effectiveness of PET using ^{18}F-fluorodeoxyglucose (^{18}FDG).[10]

Methods

The medical records of all patients with diagnosed or suspected NSCLC referred by their family physicians to the pulmonologists of the academic hospital VU University Medical Center, Amsterdam (VUMC) and the community hospital Medical Centre Alkmaar (MCA) between January 1993 and January 1995 were reviewed. Patients were identified by cross-linking databases of the Dutch Cancer Registry, the Pathological Anatomical National Register (PALGA), local surgery records and the minutes from regular multidisciplinary rounds.

The following information was extracted: demographic data, pathological data, the number and type of diagnostic investigations (excluding laboratory and lung function tests), the duration of the diagnostic process until definitive clinical TNM staging, postoperative TNM classification and follow up data. Imaging tests, punctures/biopsies (not requiring mediastinoscopy, thoracotomy, video assisted thoracoscopy, or rigid bronchoscopy) and flexible bronchoscopy were classified as non-invasive investigations. Surgical staging procedures (such as mediastinoscopy, video-assisted thoracoscopy) and rigid bronchoscopy were classified as invasive investigations.

Surgery was considered futile in cases of benign lesions, T4 lesions, macroscopic mediastinal lymph node involvement, or pleural metastasis. In addition, patients in whom a pneumonectomy was necessary to perform a complete resection but whose poor lung function allowed for only a less extended resection (such as (bi-) lobectomy) were also considered to have undergone futile surgery. Finally the diagnosis of distant metastases or local relapse during the 12 months after surgery with curative intent were also regarded as futile surgery.

Cost methodology

A cost analysis was performed to assess the costs of the staging procedures in both hospitals, including that of surgical resection and the resulting number of hospital admission days. Costs of any additional treatment were not included. The total cost of the diagnostic strategies was obtained by multiplying the number procedures performed by their individual costs. Detailed price calculations were made in the two hospitals to estimate unit costs accurately. This included the costs of personnel, materials, equipment, and overheads.

Analysis of data

Statistical analyses were performed using Student's t-test, Fischer's exact test and the Mann-Whitney U test.

Results

Three hundred and ninety five patients with suspected NSCLC were identified; 58 were medically inoperable because of severe comorbidity, leaving 337 patients (220 and 117 from the community hospital and university hospital, respectively) for further analysis; 271 (80.5%) were men, and the mean age was 64 years (range: 27-88). The age and sex distribution were comparable in the two institutions.

Staging procedures

Chest radiography and flexible bronchoscopy were performed in all patients. A CT scan of the chest was performed in 315 patients. In 17 patients metastatic disease was found prior to CT scanning. In one patient the chest lesion proved to be a metastasis of urothelial carcinoma and further diagnostic tests were directed in search of the primary tumour. Four patients refused a chest CT scan. Between the first visit and the final clinical stage classification patients underwent a mean (SD) of 5.1 (1.5) diagnostic tests, with the number in the academic hospital (5.5 (1.7)) being higher than in the general hospital (4.8 (1.5); p=0.001). However, there was no significant difference in the number of diagnostic tests between

Table 1. Staging procedures aiming at nodal or distant metastases

IMAGING		
Chest CT (+ upper abdomen)	143 (99%)	172 (89%)
Ultrasound (abdomen)	57 (40%)	87 (45%)
Bone Scintigraphy	45 (31%)	79 (41%)
CT /MRI brain	13 (9%)	41 (21%)
CT abdomen	7 (5%)	15 (8%)
INVASIVE PROCEDURES		
Mediastinoscopy	67 (47%)	41 (21%)
Other	40 (28%)	117 (61%)

patients considered operable and those deemed to be inoperable (5.0 (1.6) and 5.1 (1.6), respectively).

Apart from chest (and upper abdomen) CT scans, investigations carried out to find distant metastases in decreasing order of frequency (Table 1) were: ultrasound and abdominal CT scan (48%), bone scan (36%) and CT/MRI of the brain (16%). Further diagnostic tests were necessary to evaluate suspicious lesions in 116 of the 337 patients (34%) which resulted in an additional 112 imaging procedures and 71 biopsies (excluding invasive mediastinal procedures). Bone scintigraphy accounted for 49% of the additional imaging tests. In 46% of the 68 biopsied patients from whom biopsy specimens were taken, suspicious lesion proved to be benign and thus false positive. Bone scanning had the highest proportion of inconclusive results (36%) as compared to CT scanning of the abdomen/brain (4%) and ultrasound (3%). Twenty two (49%) of the 45 patients with inconclusive bone scans underwent surgery with curative intent.

Invasive procedures aimed at nodal and primary tumour staging were performed in 39% of all patients. In 65% of these investigations, no malignant tissue was sampled. Mediastinoscopy was performed more often in the community hospital (59/99 patients, 60%) than in the academic hospital (9/69, 13%). In the latter, mediastinoscopy was performed only in case of enlarged (>1 cm) lymph nodes on chest CT scanning.

Overall, 50% (169/337) of the patients proved to be ineligible for surgery with curative intent as a result of clinical staging.

Duration of staging and costs

The frequency distribution of the duration required for clinical staging was positively skewed towards longer durations (Fig 1). In 50% of the patients the diagnostic work up lasted more than 3 weeks (median 20 days (IQR 10-31)) where IQR= the numerical difference between the 25th and 75th centiles. This was especially true for patients eventually deemed to be operable (median 25 days (IQR 16-34) v 14 days (IQR 8-26) for clinically inoperable patients,

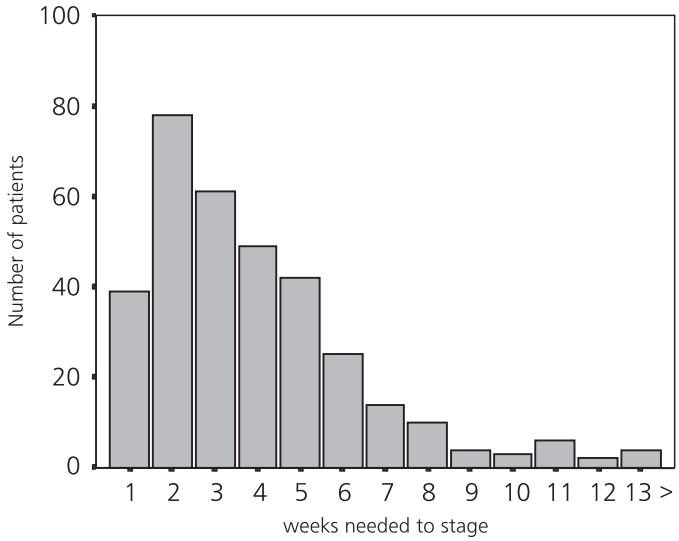


Figure 1. Time between first visit and finalising clinical stage

Table 2. Mean costs (euro) per patient (1 euro = DFL 2.20)

	VUMC	MCA
Staging (euro)	1284	3064
Operated patients	2056	3180
No surgery	498	2990
Thoracotomy*	6113	9018

VUMC = VU University Medical Centre, Amsterdam; MCA = Medical Centre Alkmaar.
* including hospital stay for staging and postoperative care, respectively.

p<0.001). The majority of patients were admitted to hospital for staging (academic hospital 61%, community hospital 95%). The mean hospital stay duration was approximately 1 week (median 9 days (IQR 2-15) for operable patients, 8 days (IQR 3-14) for inoperable patients). Hospital stay for staging and for postoperative care accounted for the majority of the costs (Table 2). The median postoperative hospital stay was 11 days in the academic hospital (IQR 9-16) compared with 15 days in the community hospital (IQR 12-17). Overall, the costs for staging were about one third that of the costs associated with thoracotomy. Higher costs were noted in the community hospital related to more hospital admissions for staging and the performance of more mediastinoscopies. Costs were similar in patients undergoing futile or non-futile surgery.

Surgery and follow-up

Planned thoracotomy was cancelled in 24/168 patients. In three the suspected lesion proved to be benign (sarcoidosis (n=2), pneumonia (n=1)), six refused thoracotomy, and the remainder were deemed to be medically unfit for surgery either because of co-morbidity (CVA (n=3),

cardiac events (n=6)) or deterioration of performance status (n=6). The exact time interval between the finalisation of clinical staging and the occurrence of the event precluding the planned thoracotomy could not be extracted from the medical records. Overall, in operated patients the median time interval between the finalisation of clinical staging and thoracotomy was 7 days (IQR 0-17). According to our criteria, surgery was futile in 74 of the remaining 144 patients (51%, 95% CI 43 to 60%; Table 3). Intraoperative staging consisted of lymph node sampling in the community hospital and mediastinal lymph node dissection in the academic hospital. Twenty patients underwent exploratory thoracotomy due to T4 lesions. Of these two patients had brachial plexus involvement, seven had direct extension into the heart and/or main stem of the pulmonary artery, six had tumours invading mediastinal structures, and five patients could not tolerate the pneumonectomy which at surgery proved to be necessary to achieve a complete resection due to poor lung function. Eleven patients had gross mediastinal lymph node involvement which had not been detected during the preoperative work up. In two patients thoracotomy was also considered explorative due to pleural metastasis. There was no significant correlation between the time between clinical and intraoperative staging for patients in whom the tumour was irresectable (median 34 days, IQR 28-49) and those with resectable tumours (median 31 days, IQR 23-45). Nineteen patients proved to have benign lesions (nine hamartomas, seven reactive lesions, two fibrosic lesions, one cyst).

Table 3. Yield of clinical staging

	MCA (n=253)		VUMC (n=142)	
After 1st screening (n)	220		117	
Clinically unresectable (III/IV)	121		48	
Clinically resectable(I/II)	99		69	
Thoracotomy	85	100%	59	100%
Benign	9	11%	10*	17%
Irresectable at surgery	10	12%	23	39%
Due to T stage	7		13*	
N stage	3		8*	
M stage	0		2	
Recurrence after radical surgery in NSCLC	13	15%	7	12%
Site: Local	0		1	
Brain	2		1	
Bone	4		0	
Liver	1		0	
Mediastinum	2		2	
Other	4		3	
Recurrence < 1yr after surgery in other malignancy	2	2%	0	
Total futile surgery	34	40%	40**	68%

VUMC = VU University Medical Centre, Amsterdam; MCA = Medical Centre Alkmaar; *P<0.01; **P=0.001

Tumour histology other than NSCLC was found in seven patients (two metastases of other primary tumours, three carcinoid tumours, one SCLC, and one mesothelioma). Three patients had bronchoalveolar cell carcinoma.
Within 12 months after apparently curative surgery 22 patients were found to have recurrent disease: symptomatic distant metastases in 14, lymph node metastases in four, local recurrence in one, and "clinically evident" relapse (as reported by the general practitioner) in three patients (Table 3). Preoperatively, 10 of the 14 patients with a distant relapse (71%) had undergone dissemination tests aimed at the affected organ. In the seven patients with tumour histology other than NSCLC, two patients with metastases of other primary tumours (colon, embryonic cell cancer) also proved to have metastases within 1 year after surgery. The other five patients remained disease-free during the follow up period and were not therefore considered to have undergone futile surgery. Five of the 144 patients who underwent curative surgery died within 3 months, one patient died between 3-12 months without evidence of disease, and six patients were lost to follow-up.
Of all the patients with NSCLC, 21% (95% CI 16 to 25%) were alive and clinically disease free at 12 months follow-up. This was not significantly different for the two hospitals (VUMC 18% (95% CI 11 to 25%), MCA 23% (95% CI 17 to 28%).

Discussion

In this systematic retrospective study staging of patients with NSCLC proved to be an intensive and often protracted effort with a disappointingly high proportion of futile thoracotomies. However, this failure rate is comparable to that reported by others.[11] VUMC is a tertiary referral centre for thoracic oncology so more patients with advanced tumours were considered for resection at this hospital. Unlike MCA (a community hospital), VUMC has a policy to staging central T3 and T4 lesions intraoperatively.
Even though the clinicians attempts to minimise the delay due to staging were successful in many patients, a considerable subset went through a more protracted process. Tumour negative test results were abundant, which is obviously useful if they alter the pretest perception of curability. Unfortunately, most procedures were simply the sequel of earlier imaging tests suggesting metastases. This reflects specificity problems of these techniques and, in the case of mediastinoscopy, also the limited sensitivity of chest CT scans. It is possible that clinicians may have ordered dissemination tests at the same time as the initial work up tests in an effort to maintain momentum in the diagnostic process. For example it is unclear what the impact of chest CT scan might have been in patients already diagnosed with brain metastases on MRI. However, in 77% of the cases dissemination tests were performed in compliance with the 1994 guidelines of Goldstraw et al.[3] (69% VUMC, 81% MCA). More imaging tests than are indicated by this guideline were performed in 15% (26% VUMC, 9%

MCA) and fewer than recommended were performed in 8% (4 % VUMC, 10% MCA). With respect to mediastinal staging procedures, practices at both hospitals were compatible with prevailing guidelines.[3,4] Even though differences in peroperative and postoperative failure patterns between hospitals appeared to reflect the differences in staging practice, the final result of diagnosis and treatment in terms of disease free 1 year survival between the two hospitals was identical.

From an economic perspective, the cost of the diagnostic tests themselves was only part of the problem. In the Netherlands costs mainly result from hospital admission for the staging procedures rather than from the tests themselves. This practice of hospitalisation is mainly intended to keep the staging process within acceptable time limits for patients. However, surveys carried out before this study (unpublished data) have shown that clinicians were under the impression that staging lasted a maximum of 2 weeks. The measurement of the actual time frame in this study confirms that questionnaires probably would have provided incorrect data in this respect.

Whole body PET allows for staging of the entire patient with a single scan, with the exception of assessment of local tumour infiltration and brain metastases. Whether its application in daily clinical practice will simplify the staging process and improve the final selection of patients for appropriate treatment is currently under study in clinical trials. There is evidence to suggest that mediastinoscopy can be omitted in cases with a negative PET scan of the mediastinum (with the exception of central tumours).[12,13]

In summary, many patients with suspected NSCLC undergo extensive investigations for clinical staging with ultimately disappointing results which apply to every level of the TNM system. Whether the current multistep process can be improved in terms of cost effectiveness by comprehensive techniques such as PET scanning remains to be shown in clinical trials. This study provides important baseline data for such trials.

References

1. Miller JD, Gorenstein LA, Patterson GA. Staging: the key to rational management of lung cancer. Ann Thorac Surg 1992;53:170-8.
2. Anonymous. Clinical practice guidelines for the treatment of unresectable non-small-cell lung cancer. Adopted on May 16, 1997 by the American Society of Clinical Oncology. J Clin Oncol 1997;15:2996-3018MH.
3. Goldstraw P, Rocmans P, Ball D et al. Pretreatment minimal staging for non-small cell lung cancer: an updated consensus report. Lung Cancer 1994;11:S1-S4.
4. van Zandwijk N. [Consensus conference on the diagnosis of lung carcinoma (in Dutch). Ned Tijdschr Geneeskd 1991;135:1915-9.
5. Dwamena BA, Sonnad SS, Angobaldo JO et al. Metastases from non-small cell lung cancer: mediastinal staging in the 1990s–meta-analytic comparison of PET and CT. Radiology 1999;213:530-6.
6. Tugwell P, Bennett KJ, Sackett DL et al. The measurement iterative loop: a framework for the critical appraisal of need, benefits and costs of health interventions. J Chronic Dis 1985;38(4):339-51.
7. Grimshaw JM,Russell IT. Effect of clinical guidelines on medical practice: a systematic review of rigorous evaluations. Lancet 1993;342:1317-22.
8. Tsang GM,Watson DC. The practice of cardiothoracic surgeons in the perioperative staging of non-small cell lung cancer. Thorax 1992;47:3-5.
9. Fergusson RJAU, Gregor AAU, Dodds RAU et al. Management of Lung cancer in South East Scotland. Thorax 1996;51:569-74.
10. Van Tinteren H, Hoekstra OS, Smit EF, and on behalf of the IKA-PLUS Study Group. Towards less Futile Surgery in Non Small Cell Lung Cancer? A Randomized Clinical Trial to Evaluate the Cost-effectivensess of Positron Emission Tomography. Controlled Clinical Trials 2001;22:89-98.
11. Lacasse Y, Bucher HC, Wong E et al. "Incomplete resection" in non-small cell lung cancer: need for a new definition. Canadian Lung Oncology Group. Ann Thorac Surg 1998;65:220-6.
12. Pieterman RM, van Putten JW, Meuzelaar JJ, et al. Preoperative staging of non-small-cell lung cancer with positron- emission tomography. N Engl J Med. 2000; 343: 254-261.
13. Vansteenkiste JF, Stroobants SG, De Leyn PR, et al. Lymph node staging in non-small-cell lung cancer with FDG-PET scan: a prospective study on 690 lymph node stations from 68 patients. Potential use of FDG-PET scan after induction chemotherapy in surgically staged IIIa-N2 non-small-cell lung cancer: a prospective pilot study. The Leuven Lung Cancer Group. J Clin Oncol 1998; 16: 2142-2149.

Chapter 3

The performance of ^{18}F-fluorodeoxyglucose positron emission tomography in small solitary pulmonary nodules

Gerarda JM Herder
Richard P Golding
Otto S Hoekstra
Emile FI Comans
Gerrit JJ Teule
Pieter E Postmus
Egbert F Smit

European Journal of Nuclear Medicine and Molecular Imaging
2004;9:1231-1236

Abstract

Background: Solitary pulmonary nodule (SPN, intraparenchymal lung mass < 3 cm) is often a diagnostic challenge. This study was performed to evaluate the diagnostic accuracy of ^{18}F-fluorodeoxyglucose positron emission tomography (^{18}FDG PET) in radiologically indeterminate SPN ≤10 mm on spiral CT.

Methods: Between August 1997 and March 2001, we identified all patients with radiologically indeterminate SPNs ≤ 10 mm who were referred for ^{18}FDG PET imaging to the VU University Medical Centre. All PET scans were retrospectively reviewed by an experienced nuclear medicine physician. PET was considered positive in cases with at least moderately enhanced focal uptake, and otherwise as negative. Lesions were considered benign on the basis of histology, no growth during 1.5 yrs or disappearance within at least 6 months.

Results: Thirty-five patients with 36 SPN ≤ 10 mm in diameter at clinical presentation were identified (one patient had two metachronous lesions). In 13 of 14 malignant nodules and in two of 22 benign nodules, diagnosis was confirmed by histology. Prevalence of malignancy was 39%. PET imaging correctly identified 30 of 36 small lesions. One lesion proved to be false negative at PET (CT:10 mm), and in five lesions, PET scans proved to be false positive. Specificity was 77% (17/22; 95% CI: 0.55-0.92), sensitivity 93% (13/14; 95% CI: 0.66-1.0), positive predictive value 72% (13/18; 95% CI: 0.46-0.90) and negative predictive value 94% (17/18; 95% CI: 0.73-1.0).

Conclusion: This retrospective study suggests that ^{18}FDG PET imaging could be a useful tool in differentiating benign from malignant SPNs ≤ 10 mm in diameter at clinical presentation. Such results may help to design larger prospective trials with structured clinical work-up.

Introduction

The availability of various diagnostic algorithms [1-3] indicates that standard clinical strategy for solitary pulmonary nodules (SPN) has not been defined. Approximately half of the patients undergoing surgical biopsy of an indeterminate lung nodule proved to have benign disease.[4-7] Among indeterminate lesions smaller than 10 mm, the prevalence of a benign lesion could be even higher than among those lesions > 10 mm.[8]

With the increased interest in the use of low-dose spiral computed tomography (CT) for early lung cancer detection, the number of coincidental SPNs will increase. Furthermore, repeated screening and technical improvements may result in the detection of very small lesions (2-10 mm in diameter) for which it is even more difficult to establish a definitive diagnosis. The yield of flexible fibre-optic bronchoscopy (20-62%) is directly related to the visibility and the size of the lesions [9,10]; similarly, the accuracy of transthoracic needle biopsy depends on the size of the lesion and ranges from 74 to 96%.[11,12] The chance of a non-diagnostic result is significant in lesions smaller than 2 cm, and the risk of a pneumothorax is substantial (3.1-41.7% [13]). Since a wait-and-see policy carries potentially adverse effects on outcome, there is a demand for accurate non-invasive tests to prevent surgical interventions for benign lesions.

It is important to identify malignant nodules as early as possible because 5-year survival in patients with resected non-small cell lung cancer (NSCLC) stage IA that has been resected can be 80%.[14,15] In patients with proven or strongly suspected NSCLC, adding ^{18}F-fluorodeoxyglucose positron emission tomography (^{18}FDG PET) to the diagnostic work-up improves the selection of surgical candidates [16], but its role in radiologically indeterminate SPNs is less well established.[17] The published evidence on the accuracy of ^{18}FDG PET in characterising SPNs consists of over 1,400 published cases [18], in which the mean sensitivity and specificity were 96% and 73.5%, respectively. However, it was noted that different criteria are used to classify the PET results. Standardised uptake values (SUVs) may provide numerical thresholds to differentiate malignant and benign lesions, but visual assessment may be at least as sufficient [19] and is highly reproducible.[20] However, in small lesions it is not yet clear which level of tracer uptake in the lesion relative to background activity results in both the best sensitivity and specificity. The lesion contrast on ^{18}FDG PET will decrease in lesions with a diameter smaller than twice the spatial resolution of the system, which is typically 5-7 mm for a state-of-the-art BGO full ring scanner.

The aim of the present study was to assess the optimal operating characteristics of visual assessment, and to provide an initial estimate of the diagnostic accuracy of ^{18}FDG PET in SPNs with a maximum diameter of 10 mm.

Materials and methods

Between August 1997 and March 2001 all patients with an SPN who underwent PET imaging were retrospectively identified from the database of the PET centre at the VU University Medical Centre (VUmc). Patients are registered in the database of the clinical PET centre according the American College of Radiology Index for Radiological Diagnoses. Two hundred and twenty-two patients with pulmonary lesions ≤ 30 mm were identified using the above mentioned search strategy. In 199 of these patients a chest CT was obtainable. Thirty-four patients were not included because of a radiographically occult lung cancer or multiple lung lesions. In two patients no definitive diagnosis (malignant or benign) could be obtained by pathology or follow-up due to early death. All patients with an SPN ≤ 10 mm in diameter on CT were eligible for the present study. An independent experienced radiologist (R.P.G.) reviewed all CT scans, blinded for clinical pretest data, ^{18}FDG PET result and outcome. If the SPN was ≤ 10 mm in diameter and without typically benign calcifications, it was classified as indeterminate and included in the study. Prior malignancy and diabetes were not exclusion criteria.

CT and ^{18}FDG PET

We reviewed chest CT scans with the shortest interval between PET and CT. Spiral CT was performed with axial slice thickness of 10 mm; in one patient axial slice thickness of CT was 4 mm and in another patient, 5 mm. In three patients high-resolution CT (1 mm axial slice thickness) was performed. Intravenous contrast was used in some but not all institutions. Images were analysed with both a mediastinal and a pulmonary parenchymal setting. The parenchymal setting was used to measure lesion diameter.

PET was performed with a dedicated full ring BGO scanner (ECAT EXACT HR+, CTI/Siemens). Emission scans, were acquired in 2D mode (5-7 min/bed position), approximately 60 minutes after intravenous injection of 370 MBq (10mCi) ^{18}FDG. Patients were asked to fast for at least 6 h prior to PET. All scans were corrected for decay, scatter and randoms. Scans were reconstructed using ordered subset expectation maximisation (OSEM) with two iterations and 16 subsets, followed by post-smoothing of the reconstructed image using a Hanning 0.5 filter resulting in a transaxial spatial resolution of 7 mm at full-width half-maximum.

Data analysis

A visual analysis of ^{18}FDG PET scan was performed by an experienced nuclear medicine physician who was blinded to patient outcome. In a first session, he was blinded to all clinical information other than the nature of the study. In the second session, localisation and lesion diameter were also provided. To simulate usual reporting practice, ^{18}FDG PET results using pre-test data were used for analysis. Intensity of FDG uptake of pulmonary lesions was visually associated with mediastinal background activity [21] and scored using a four-point

scale (absent, faint, moderate or intense). Final classification was based on histopathological findings or clinical and radiological follow-up. Radiological follow-up typically consisted of repeat CT scan of the thorax. Lesions were considered malignant on the basis of pathology or growth at radiological follow-up. Lesions were classified as benign on the basis of pathological findings, disappearance of the lesion at radiological follow-up, or absence of growth within an observation period of at least 1.5 years.

Probability of malignancy

If applicable, we assessed the pre-PET probability of malignancy, using the model according to Swensen [22], which applies to patients with indeterminate SPN without cancer within the past 5 years or a history of primary lung cancer.

This clinical prediction model for malignancy in SPN expresses the probability of malignancy as a function of six variables, three clinical (age, current or former cigarette smoker, history of cancer more than 5 years ago) and three radiographic (diameter, spiculation and location in upper lobe).

Clinical impact of ^{18}FDG PET

Between July 1997 and July 2001, the impact of ^{18}FDG PET on diagnostic understanding and management was prospectively assessed using questionnaires. These forms were completed by the referring physicians and included information on the intended treatment plan prior to ^{18}FDG PET, the actual therapy choice after ^{18}FDG PET, and a post hoc clinical assessment of the impact of ^{18}FDG PET on diagnostic understanding and management.[23,24] In the first questionnaire, information was requested regarding the histological diagnosis (if known), a definition of the current diagnostic problem, a differential diagnostic consideration, the results of diagnostic tests already performed and any planned diagnostic tests. In addition, the referring physician was requested to outline the intended patient management plan if ^{18}FDG PET scanning was not available. The second questionnaire requested information regarding the working diagnosis and planned treatment after ^{18}FDG PET scanning in addition to any diagnostic tests that had been ordered as a direct consequence of the ^{18}FDG PET scan result. In the final questionnaire, the referring physician was requested to convey the final diagnosis and to rate the overall usefulness of ^{18}FDG PET separately in terms of diagnostic understanding and therapy choice according to the method of Wittenberg et al.[24] This method involves using a five-point ordinal scale, with higher scores representing increasing positive impact.

Statistical analysis

Sensitivity, specificity, accuracy, positive predictive value and negative predictive value of ^{18}FDG PET were determined. For each parameter the 95% confidence intervals (95% CI) were calculated using confidence interval analysis version 1.0. Receiver operating characteristic

(ROC) curves for clinical prediction model and PET were derived using SPSS 10.0 and were evaluated by comparing the areas under the ROC curve.

Table 1. Clinical and histological findings

Lesion	Size on CT (mm)	History of cancer < or > 5 years	Location	Intensity of FDG uptake	Pathology/ FU (growth/no growth/ disappearance)	Time between PET and CT (days)
1	3		RLL	Moderate	Lymphoma	29
2	5	Yes<5y	RUL	Moderate	NSCLC	0
3	5	Yes<5y	LLL	Intense	Lymphoma	71
4	5		RML	Absent	FU (disappearance)	42
5	5		RUL	Faint	FU (no growth)	25
6	7		LUL	Moderate	Sarcoidosis	3
7	7		RML	Absent	FU (disappearance)	8
8	8		LUL	Intense	NSCLC	33
9	10	Yes<5y	LUL	Intense	Metastasis	55
10	10		RLL	Moderate	NSCLC	0
11	10		LUL	Faint	Carcinoid	0
12	10	Yes >5y	RUL	Intense	NSCLC	99
13	10		LUL	Moderate	NSCLC	93
14	10		RUL	Intense	Carcinoid	55
15	10		LLL	Moderate	SCLC	27
16	10		LLL	Intense	NSCLC	24
17	10		RUL	Intense	NSCLC	73
18	10		RUL	Intense	Fibrosis/granulomatosis	16
19	10		RLL	Faint	FU (disappearance)	11
20	10		RUL	Intense	FU (no growth)	20
21	10		LUL	Absent	FU (no growth)	14
22	10		RUL	Absent	FU (no growth)	29
23	10	Yes >5y	LUL	Absent	FU (no growth)	40
24	10		LUL	Moderate	FU (growth)	65
25	10		RML	Moderate	FU (disappearance)	45
26	10		RUL	Faint	FU (no growth)	6
27	10	Yes<5y	RML	Absent	FU (no growth)	66
28	10		LUL	Moderate	FU (disappearance)	17
29	10		RML	Absent	FU (no growth)	16
30	10		RML	Absent	FU (disappearance)	36
31	10		LUL	Absent	FU (no growth)	62
32	10		RML	Absent	FU (disappearance)	42
33	10		LUL	Absent	FU (no growth)	60
34	10		LLL	Absent	FU (no growth)	41
35	10		LUL	Absent	FU (disappearance)	71
36	10		RUL	Absent	FU (disappearance)	2

FU, Follow-up; LLL, left lower lung; LUL, left upper lung; RLL, right lower lung; RML, right middle lung; RUL, right upper lung

Results

In 35 patients (36 pulmonary nodules) the pulmonary nodule was ≤ 10 mm in diameter (one patient had two separate lesions). Referring physicians were pulmonologists from university (n=13) and community (n=22) hospitals. Fifty-seven percent of the patients were female. Patients' mean age was 61 years (SD 10), with 13 pack years of smoking (median; range 0-55). Six patients had a history of prior malignancy. Clinical and histological data are listed in Table 1. In all patients SPN was located peripherally. In the 32 cases to which the model of Swensen was applicable (no prior cancer within 5 yrs), the mean pretest probability of malignancy was 15% (SD 12). The distribution of SPNs was: left lung 16, upper lobes 22, lower lobes 7 and middle lobes 7. The mean size at CT was 9 mm (range 3-10), and 28 measured 10 mm.
Fourteen of 36 nodules proved to be malignant (prevalence 39%). The final diagnosis was confirmed by histology in 13/14 malignant (NSCLC 7, lymphoma 2, SCLC 1, carcinoid 2, metastasis 1) and in two of 22 benign nodules (granuloma, sarcoidosis). In the remaining lesions, the median duration of radiological follow-up was 293 days for lesions that disappeared [IQR 119-429 (IQR, inter quartile range, i.e. the numerical difference between the 25th and 75th centiles), n=10], and 726 days in lesions without growth (IQR 564-1038, n=11). Radiological progression was found in a single patient who was inoperable due to poor pulmonary function. In eight patients CT was performed after ^{18}FDG PET. In the remaining 28 patients median time between CT and ^{18}FDG PET scan was 35 days (IQR 16-59). In ten patients, thoracotomy followed PET after a median of 57 days (IQR 27-119).

^{18}FDG PET results

^{18}FDG PET readings with and without knowledge of localisation and size were identical in all but one case, in which the lesion was only recognised at ^{18}FDG PET after this additional information had been made available. This lesion was then classified as a positive ^{18}FDG PET result (moderate uptake). For further analysis, we used the readings that included knowledge of size and localisation since this situation closely resembles clinical practice. ROC analysis revealed that using a at least moderately enhanced uptake as a cut-off for ^{18}FDG PET test positivity (i.e. suspicious for malignancy) yielded a sensitivity of 93% (13/14; 95% CI: 0.66-1.0) and a specificity of 77% (17/22; 95% CI: 0.55-0.92) (Fig 1). Using this threshold, ^{18}FDG PET imaging correctly identified 30 of 36 SPNs; it was false negative in a 10-mm lesion which proved to be carcinoid (primary) at surgery. This patient had a fasting state blood glucose of 6 mmol/l. The referring physician decided to proceed to thoracotomy to obtain histological diagnosis of the SPN. ^{18}FDG PET was false positive in five cases: two proved to be fibrosis and sarcoidosis at pathology, in another two cases the lesion disappeared and in the final case the lesion was stable at prolonged follow-up.

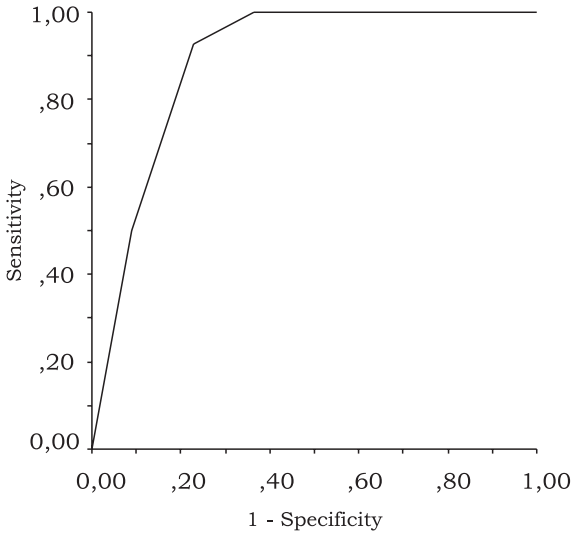


Figure 1. ROC curve for ^{18}FDG PET results. Sensitivity (true positive rate) vs 1-specificity (false positive rate)

Table 2. Predicted negative and positive predictive values as a function of prevalence of malignancy, using test performance characteristics of the present study (i.e. sensitivity 93%, specificity 77%).

	NPV	PPV
Prevalence 10%	0.99	0.31
Prevalence 39%	0.95	0.72
Prevalence 55%	0.90	0.83
Prevalence 90%	0.55	0.97

NPV, Negative predictive value; PPV, positive predictive value

Clinical impact

In 28 of 36 ^{18}FDG PET scans information was available on the impact on diagnostic understanding and management. Patients included in ongoing prospective studies (n=6) and patients with incomplete forms from referring physician (n=2) were not included in the assessment of this impact (i.e. the response rate was 93%). ^{18}FDG PET had a positive influence on diagnostic understanding in 22 of 28 (79%) cases. According to referring physicians ^{18}FDG PET resulted in beneficial change of therapy in 16 of 28 (57%), most importantly obviating the need for surgery in 12 of 16 patients. In three other patients, surgery was substituted for observation. A management change within surgery occurred in one patient.

Discussion

In this retrospective study, a simple visual assessment of ^{18}FDG PET scans adequately classified radiologically indeterminate SPNs ≤10 mm in diameter at clinical presentation. In fact, our

estimates of accuracy fit well into the summary ROC curve provided by Gould et al [18], which was predominantly based on studies of larger SPNs.

A visual analysis of ^{18}FDG PET scan was performed because it has been shown that SUV methodology and implementation is less straightforward than was often assumed. There is still debate about the appropriate normalisations to be used, and more recently it was demonstrated that results of SUVs strongly depend on image reconstruction methodology.[25] Moreover, a recent meta-analysis failed to show that semi-quantitative image interpretation improves the accuracy of ^{18}FDG PET.[18] Finally, the lack of standardisation in the current PET literature and practice strongly compromises the theoretical advantage of "objective" measurements.

In practice, management of SPNs is based on the perceived probability of malignancy (depending on radiological and clinical variables). An important reason for variation in management lies in discrepancies between clinicians in estimating the probability of malignancy (especially in the case of intermediate probabilities).[26] To reduce the effect of such heterogeneity (i.e. inclusion of lesions obviously malignant or benign at CT scanning) in this study, an experienced radiologist identified the patients with truly radiologically indeterminate SPNs.

Every imaging technique has a detection limit. In nuclear medicine technology, the dominant factor of lesion visualisation is the relative tracer uptake versus the background. For FDG-avid tumours like melanoma, it has been shown that accuracy clearly declines below the spatial resolution of the present generation of scanners.[27] Until respiratory gating techniques are generally available, the ^{18}FDG signal arising from peripheral lung lesions will be smeared due to the superficial breathing during acquisition, which typically lasts for 5 minutes with the present generation of scanners. Within the lung, the background level of ^{18}FDG activity varies, according to both ventrodorsal and craniocaudal gradients, being lowest in the upper and anterior lungs.[28] Finally, just above the diaphragm, scatter from photons arising from the liver may adversely affect both image quality and lesion detectability. In our study, these factors did not affect the accuracy of PET, but the proportion of lower lobe lesions was relatively small (7/36). Based on the experience with other tumours, we would expect that the negative predictive value of currently available PET camera's for peripheral lung lesions < 5-7 mm may be insufficient for clinical implementation.

Thirty-nine percent of the SPNs in our study was malignant. This is clearly lower than the reported prevalence in many other studies (ranging from 55-88% [18]), but clearly higher than the prevalence of malignancy in SPNs detected in lung cancer screening programs (range 7%-24% [29,30]). This suggests that the negative predictive value of 94% observed in the present study may not be applicable in such conditions (Table 2). Similarly, depending on the specific clinical situation, one might consider other thresholds for test positivity with ^{18}FDG PET.

A limitation of the retrospective nature of the present study is the median time lapse of 35 days between spiral CT and ^{18}FDG PET. Presuming a high tumour doubling time of 30 days, exponential growth and a spherical nodule, the expected diameter of a malignant pulmonary nodule with an initial diameter of 5 mm would be 6.2 mm after 28 days, and 12.4 mm for

lesions of initially 10 mm diameter.[31] However, retrospective calculations (starting with size at pathology) for our patients using a high tumor doubling time of 30 days, resulted in tumour sizes within the 10 mm limit at time of ^{18}FDG PET.

Conclusion

In summary, together with the clinically perceived yield of adding ^{18}FDG PET to standard probability estimation, our accuracy data suggest that studies are warranted on the accuracy and clinical utility of ^{18}FDG PET in small indeterminate SPNs detected at CT. Ultimately, the focus of such investigations will have to extend beyond accuracy to include analysis of relevant patient outcome measures and costs.

References

1. Cummings SR, Lillington GA, Richard RJ. Managing solitary pulmonary nodules. The choice of strategy is a "close call". Am Rev Respir Dis 1986;134:453-60.
2. Ost D, Fein A. Evaluation and management of the solitary pulmonary nodule. Am J Respir Crit Care Med 2000;162:782-87.
3. Swensen SJ, Jett JR, Payne WS, et al. An integrated approach to evaluation of the solitary pulmonary nodule. Mayo Clin Proc 1990;65:173-86.
4. Bernard A. Resection of pulmonary nodules using video-assisted thoracic surgery. The Thorax Group. Ann Thorac Surg 1996;61:202-4.
5. Keagy BA, Starek PJ, Murray GF, et al. Major pulmonary resection for suspected but unconfirmed malignancy. Ann Thorac Surg 1984;38:314-16.
6. Mack MJ, Hazelrigg SR, Landreneau RJ, et al. Thoracoscopy for the diagnosis of the indeterminate solitary pulmonary nodule. Ann Thorac Surg 1993;56:825-30.
7. Siegelman SS, Zerhouni EA, Leo FP, et al. CT of the solitary pulmonary nodule. AJR 1980;135:1-13.
8. Suzuki K, Nagai K, Yoshida J, et al et al. Video-assisted thoracoscopic surgery for small indeterminate pulmonary nodules: indications for preoperative marking. Chest 1999;115:563-68.
9. Baaklini WA, Reinoso MA, Gorin AB, et al. Diagnostic yield of fiberoptic bronchoscopy in evaluating solitary pulmonary nodules. Chest 2000;117:1049-54.
10. Hsu C. Cytologic diagnosis of lung tumors from bronchial brushings of Chinese patients in Hong Kong. Acta Cytol 1983;27:641-46.
11. Hayata Y, Oo K, Ichiba M, et al. Percutaneous pulmonary puncture for cytologic diagnosis–its diagnostic value for small peripheral pulmonary carcinoma. Acta Cytol 1973;17:469-75.
12. Li H, Boiselle PM, Shepard JO, et al. Diagnostic accuracy and safety of CT-guided percutaneous needle aspiration biopsy of the lung: comparison of small and large pulmonary nodules. AJR 1996;167:105-9.
13. Lacasse Y, Wong E, Guyatt GH, et al. Transthoracic needle aspiration biopsy for the diagnosis of localised pulmonary lesions: a meta-analysis. Thorax 1999;54:884-93.
14. Patz EF, Jr., Rossi S, Harpole DH, Jr., et al. Correlation of tumor size and survival in patients with stage IA non- small cell lung cancer. Chest 2000;117:1568-71.
15. van Rens MT, de la Riviere AB, Elbers HR, et al. Prognostic assessment of 2,361 patients who under-

went pulmonary resection for non-small cell lung cancer, stage I, II, and IIIA. Chest 2000;117:374-79.

16. van Tinteren H, Hoekstra OS, Smit EF, et al. Effectiveness of positron emission tomography in the preoperative assessment of patients with suspected non-small-cell lung cancer: the PLUS multicentre randomised trial. Lancet 2002;359:1388-93.
17. Balk E, Lau J. PET scans and technology assessment: deja vu? JAMA 2001;285:936-37.
18. Gould MK, Maclean CC, Kuschner WG, et al. Accuracy of positron emission tomography for diagnosis of pulmonary nodules and mass lesions: a meta-analysis. JAMA 2001;285:914-24.
19. Lowe VJ, Fletcher JW, Gobar L, et al. Prospective investigation of positron emission tomography in lung nodules. J Clin Oncol 1998;16:1075-84.
20. Bos R, Der Hoeven JJ, van der Wal E, et al. Biologic correlates of (18)fluorodeoxyglucose uptake in human breast cancer measured by positron emission tomography. J Clin Oncol 2002;20:379-87.
21. Vansteenkiste JF, Stroobants SG, Dupont PJ, et al. FDG-PET scan in potentially operable non-small cell lung cancer: do anatometabolic PET-CT fusion images improve the localisation of regional lymph node metastases? The Leuven Lung Cancer Group. Eur J Nucl Med1998;25:1495-501.
22. Swensen SJ, Silverstein MD, Ilstrup DM, et al. The probability of malignancy in solitary pulmonary nodules. Application to small radiologically indeterminate nodules. Arch Intern Med 1997;157:849-55.
23. Herder GJ, Tinteren H, Comans EF, et al. Prospective use of serial questionnaires to evaluate the therapeutic efficacy of 18F-fluorodeoxyglucose (FDG) Positron Emission Tomography (PET) in (suspected) lung cancer. Thorax 2003;58:47-51.
24. Wittenberg J, Fineberg HV, Black EB, et al. Clinical efficacy of computed body tomography. AJR 1978;131:5-14.
25. Boellaard, R, Krak, NC, Hoekstra, OS, et al. Effects of image acquisition, reconstruction and ROI definition on the accuracy and reproducibility of standard uptake values. J Nucl Med 2002;44: 260.
26. Baldwin DR, Eaton T, Kolbe J, et al. Management of solitary pulmonary nodules: how do thoracic computed tomography and guided fine needle biopsy influence clinical decisions? Thorax 2002;57:817-22.
27. Crippa F, Leutner M, Belli F, et al. Which kinds of lymph node metastases can FDG PET detect? A clinical study in melanoma. J Nucl Med 2000;41:1491-94.
28. Miyauchi T, Wahl RL. Regional 2-[18F]fluoro-2-deoxy-D-glucose uptake varies in normal lung. Eur J Nucl Med 1996;23:517-23.
29. Henschke CI, McCauley DI, Yankelevitz DF, et al. Early lung cancer action project: a summary of the findings on baseline screening. Oncologist 2001;6:147-52.
30. Sobue T, Moriyama N, Kaneko M, et al. Screening for lung cancer with low-dose helical computed tomography: anti-lung cancer association project. J Clin Oncol 2002;20:911-20.
31. Yankelevitz DF, Reeves AP, Kostis WJ, et al. Small pulmonary nodules: volumetrically determined growth rates based on CT evaluation. Radiology 2000;217:251-56.

Chapter 4

Clinical prediction model to characterise pulmonary nodules: validation and added value of ^{18}FDG PET

Gerarda JM Herder
Harm van Tinteren
Richard P Golding
Piet J Kostense
Emile FI Comans
Egbert F Smit
Otto S Hoekstra

Chest 2005;128:2490-2496

Abstract

Background: The added value of ^{18}FDG PET as a function of pre-test risk assessment in indeterminate pulmonary nodules is still unclear.

Objective: To obtain an external validation of the prediction model according to Swensen and colleagues, and to quantify the potential added value of ^{18}FDG PET as a function of its operating characteristics in relation to this prediction model, in a population of patients with radiologically indeterminate pulmonary nodules.

Methods: Between August 1997 and March 2001, all patients with an indeterminate solitary pulmonary nodule referred for ^{18}FDG PET were retrospectively identified from the database of the PET centre at the VUmc.

Results: One hundred six patients were eligible, and 61 (57%) proved to have malignant nodules. The goodness-of-fit statistic for the model (according to Swensen) indicated that the observed proportion of malignancies did not differ from the predicted proportion (p=0.46). ^{18}FDG PET results classified using the 4-point intensity scale reading yielded an area under the evaluated receiver operating characteristic curve of 0.88 (95% CI 0.77-0.91). The estimated difference of 0.095 (95% CI: -0.003-0.193) between the ^{18}FDG PET results classified using the 4-point intensity scale reading and the area under the curve (AUC) from the Swensen prediction was not significant (p=0.058). ^{18}FDG PET added to the predicted probability calculated by the Swensen model improves the AUC by 13.6 % (95% CI: 6-21; p=0.0003).

Conclusion: The clinical prediction model of Swensen et al. [1] was proven to have external validity. However, especially in the lower range of its estimates, the model may underestimate the actual probability of malignancy. The combination of visually read ^{18}FDG PET scans and pretest factors appears to yield the best accuracy.

Introduction

Radiologically indeterminate solitary pulmonary nodules (SPN) are a diagnostic challenge in pulmonary medicine. Currently, most SPN's are discovered by plain chest films. With the introduction of computed tomography (CT) screening for lung cancer, the number of SPN will strongly increase. Unfortunately, after a full noninvasive evaluation the diagnosis may still be unclear.

One comprehensive cost-effectiveness analysis proposed a diagnostic approach which strongly relied upon clinical risk assessment.[2] This probability estimation was based on clinical as well as radiological parameters, and has been developed and preliminarily validated in a United States population.[1] The cost-effectiveness analysis included the potential role of ^{18}F-fluorodeoxyglucose (^{18}FDG) positron emission tomography (PET) scanning. However, the criteria for judging test results with ^{18}FDG PET are heterogeneous [3], and standardisation would be desirable. The added value of ^{18}FDG PET as a function of pre-test risk assessment still needs to be established.

The aims of the present study were two-fold: first, to obtain an external validation of the prediction model; and second, to quantify the potential added value of ^{18}FDG PET as a function of its operating characteristics in relation to this prediction model in a population of patients with radiologically indeterminate pulmonary nodules.

Methods

Between August 1997 and March 2001, all patients with an indeterminate SPN, detected during normally clinical work in both university and community hospital setting, referred for ^{18}FDG PET were retrospectively identified from the database of the PET centre at the VU University Medical Centre (VUmc). In our database, characteristics of all patients are registered using a modified version of the American College of Radiology Index for Radiological Diagnoses.

An independent experienced radiologist (RPG), who was blinded to clinical pretest data, ^{18}FDG PET results and outcome reviewed all CT scans. Patients were eligible for the study if the SPN was $\leq$ 30 mm in diameter on CT and without typically benign calcifications. Patients with prior malignancies within the past 5 years before ^{18}FDG PET scanning, unknown history of malignancy or without a definitive clinical diagnosis, or patients lost to follow-up were not eligible.

All medical records were reviewed to obtain the following data: age, gender, smoking status (current or former cigarette smoker, number of pack-years), history of malignancy (date and kind of malignancy), pathology (conclusion and date) and last date of clinical and radiological follow-up (including: disappearance of the SPN or decreased size, no growth and growth).

Imaging CT

Spiral CT scanning was performed with axial slice thickness of 10 mm in 96 patients, 5 mm in 2 patients and 4 mm in another patient. In seven patients a high-resolution CT (1 mm axial slice thickness) was performed. Intravenous contrast was used at some but not all institutions.
Images were analysed with mediastinal as well as pulmonary parenchymal settings. The SPN diameter (*ie,* the mean of diameters in the transverse plane in millimeters), its location (*ie,* upper lobe or elsewhere) and the presence of spiculae (*ie,* < 50% or ≥ 50% of the circumference) were recorded.

Imaging ^{18}FDG PET

^{18}FDG PET was performed with a dedicated full-ring BGO scanner (ECAT EXACT HR+, CTI/Siemens;Knoxville, TN). Emission scans, were acquired in the two dimensional mode (5-7 min/bed position), approximately 60 minutes after intravenous injection of 370 MBq of ^{18}FDG. Patients were asked to fast for at least 6 hours prior to undergoing the PET scan. All scans were corrected for decay, scatter and randoms, and were reconstructed using ordered subset expectation maximisation (OSEM) with 2 iterations and 16 subsets followed by post smoothing of the reconstructed image using a Hanning 0.5 filter, resulting in a transaxial spatial resolution of 7 mm at full-width half-maximum.
One experienced nuclear medicine physician (EFC) reviewed all ^{18}FDG PET scans. A visual analysis of ^{18}FDG PET scan was performed blinded to patient outcome. To simulate usual reporting practice, localisation, and diameter of the lesion were provided. The intensity of ^{18}FDG uptake was scored using a 4-point scale (0, absent; 1, faint; 2, moderate; or 3, intense). Inter-observer variation of this classification system was assessed by asking seven relatively inexperienced nuclear medicine physicians, who were blinded for all clinical and radiological information other than "SPN," to score a randomly chosen subset (25% of the present material). Semiquantitative analysis was performed using tumour normal lung tissue ratio (T/N).

Diagnosis

Final classification was based on histopathological findings or clinical and radiological follow-up. Time of follow-up was defined as time between ^{18}FDG PET imaging and histological diagnosis or date of last radiological follow-up. Radiological follow-up typically consisted of repeat chest CT scans. Lesions were classified as benign in case of benign pathological findings, disappearance of the lesion at radiological follow-up, or no change in size within an observation period of at least 1 year. Lesions were considered malignant on the basis of pathology or growth at radiological follow-up.

Clinical prediction model according to Swensen [1]

This model expresses the probability of malignancy as a function of 3 clinical and 3 radiographic variables as follows:

$$\text{probability of malignancy} = 1 / (1 + e^{-x})$$

where x = -6.8272 + 0.0391{age} + 0.7917 {cigarettes} + 1.3388 {cancer} + 0.1274 {diameter} + 1.0407 {spiculation} + 0.7838 {upper}; e is the base of natural logarithms; *age* is the patient's age (years); *cigarettes* is 1 if the patient is a current or former smoker (otherwise, 0); *cancer* is 1 if the patient has a history of extrathoracic cancer diagnosed > 5 years ago (otherwise, 0); *diameter* is the diameter of the SPN in millimeters; *spiculation* is 1 if the edge of the SPN has spiculae (otherwise is 0), and *upper* is 1 if the SPN is located in an upper lobe (otherwise is 0). The model was validated for an American population with a 26.4% prevalence of malignancy.

Statistical analysis

The model fit was assessed by a goodness-of-fit for binary logistic regression, [4] as implemented by Harrell et al,[5] where high p values indicate a well-calibrated model. The predictive ability was expressed by various statistics, among others the area under the receiver operating characteristics (ROC) curve. The area under the curves (AUCs) were compared using the method described by DeLong et al [6] and logistic regression models were compared using Akaike's information criterion (AIC). First the accuracy of the prediction model of Swensen et al [1] was determined on the study population. Second, the characteristics of ^{18}FDG PET, as a univariate test with 4 categories, were calculated. Finally, the added value of ^{18}FDG PET to the Swensen model [1] was explored. A nomogram was constructed using the pretest probability of the model of Swensen et al combined with the value of ^{18}FDG PET. Interobserver variation of ^{18}FDG PET classification was analysed with intraclass correlation coefficients. Extensive use was made of programs developed (S-plus, version 6.2; Insightful; Seattle,WA) by Harrell et al.[5]

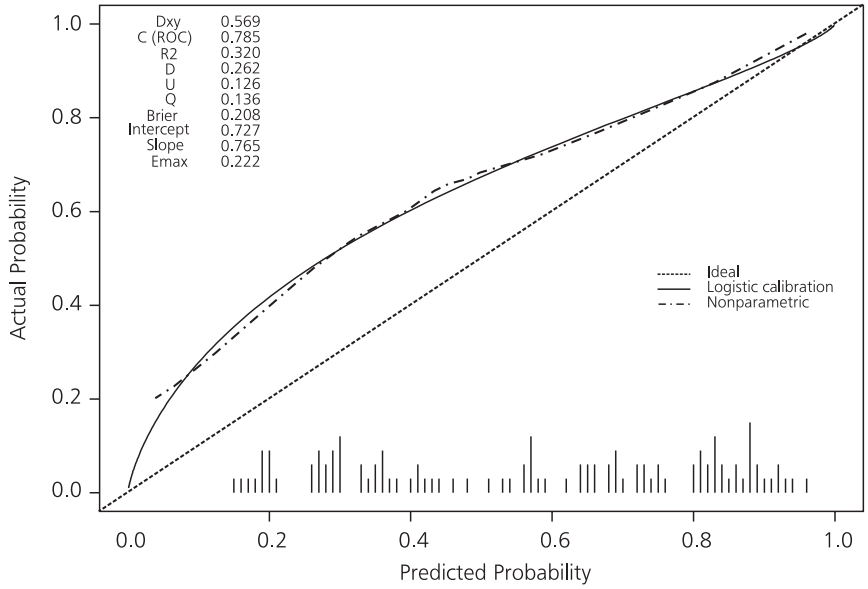
Results

In total, 106 eligible patients were identified of whom 61 (57.5%) proved to have malignant nodules. Referring physicians were pulmonologists from university (n=25) and community (n=81) hospitals. Fifty-eight percent of the patients were male and their mean age was 64 years (range 32-85) (Table 1). The diagnosis of malignancy was based on histopathological results in 55 patients and on radiological growth of the lesion in 6 patients. The diagnosis of a benign lesion was based on the stabilisation or spontaneous decrease in size of the lesion at follow-up CT in 40 patients and on the histopathological result in 5 patients. In patients with radiological stable SPNs (n=23) the median follow-up was 646 days (interquartil range 413-925 days), and only 6 had a follow-up < 365 days (with a minimum of 203 days) versus 205 days (interquartile range 143-398 days) in the 17 patients with shrinking or disappearing lesions. Interquartile range, i.e. the numerical difference between the 25th and 75th centiles.

Table 1. Baseline demographic data (n=106):

	Malignant (n=61)	Benign (n=45)
Mean age, years (sd)	66 (10)	60 (13)
Male/female	35/26	27/18
Current or former smoker	53	26
Cancer > 5 yrs ago	9	1
Spicula ≥ 50%	34	8
Location		
Upper lobe	40	30
Elsewhere	21	15
diameter (mm)		
≤ 10	11	22
11-20	26	16
21-30	24	7
^{18}FDG PET uptake		
Absent	1	26
Faint	1	6
Moderate	16	7
Intense	43	6

Figure 1. Validation of the clinical prediction model of Swensen in 106 patients with SPNs



Computed indexes and statistics: Dxy: Somers Dxy rank correlation between actual and predicted probability; C: ROC area; R^2: Nagelkerke-Cox-Snell-Maddala-Magee R-squared index; D: discrimination index (logistic model LR-chi-square-1/n; U: unreliability index (chi-square with 2d.f. for testing unreliability); Q: quality index Q, Brier score (average squared difference between actual and predicted probability); Emax: maximum absolute difference in predicted and calibrated probabilities.

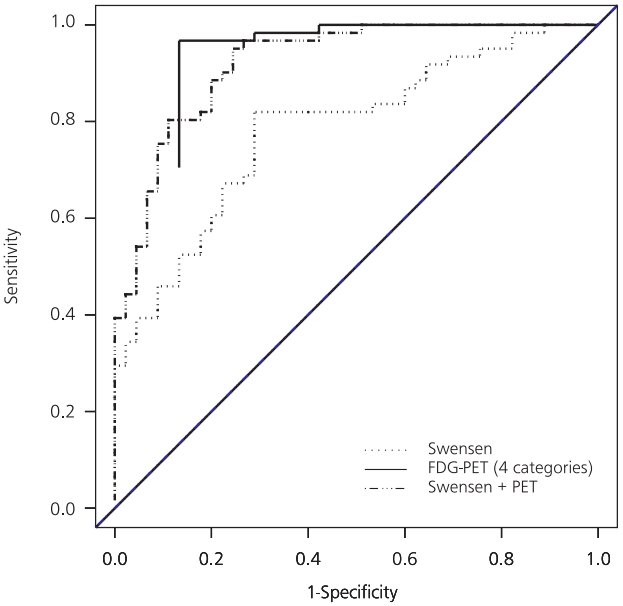
Validation of the Swensen model [1]

The goodness-of-fit statistic for the model indicated that the observed proportion of malignancies did not differ from the predicted proportion (p=0.46). The probability of malignancy was calculated using the complete model (*eg* variables with specified coefficients) of Swensen. The ROC-AUC was 0.79 (95% confidence interval [CI]: 0.70-0.87). A calibration curve of the model including a series of statistics is shown in Figure 1.

Operating characteristics of ^{18}FDG PET

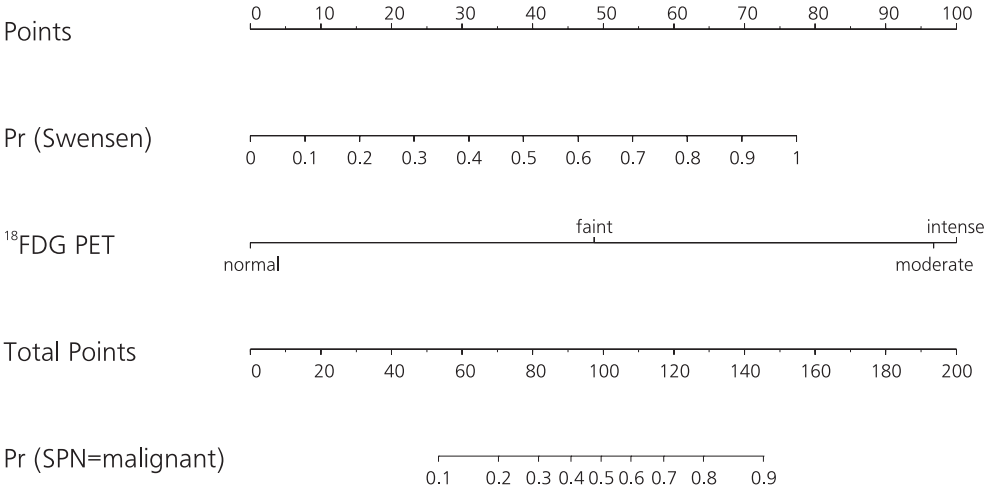
^{18}FDG PET results classified using the 4-point intensity scale reading (Figure 2) yielded an ROC-AUC value of 0.88 (95% CI 0.77-0.91). A tumour normal tissue (T/N) ratio on these data showed identical AUC-ROC values (AUC 0.87 (95%CI 0.80-0.94)). All other analyses were performed using the 4-point intensity scale reading. The estimated difference of 0.095 (95% CI: -0.003-0.193) with the AUC from prediction of the Swensen et al [1] was not significant at p=0.058. Classifying the 6.6% proportion (n=7) with faintly enhanced ^{18}FDG uptake as negative, yielded a sensitivity of 96.7% (95% CI: 87.6-99.4; 59/61), a specificity of 71.1% (95% CI: 55.5-83.2; 32/45) and an accuracy of 86% (95% CI: 77.4-91.6). Two nodules without enhanced ^{18}FDG uptake proved to be papillary adenocarcinoma (diameter, 30 mm) and carcinoid (diameter, 10 mm) at pathology. Thirteen nodules with increased ^{18}FDG uptake were classified as benign, with histological diagnoses of fibrosis (one patient), hematoma (two patients), reactive granulomatosis (two patients), radiological regression (seven patients) or no growth (one patient). Inter-observer correlation of visual analysis of ^{18}FDG PET using intensity scales was 0.87 (95% CI 0.79-0.93).

Figure 2. ROC curves for the prediction model of Swensen et al [1] and for a model combining Swensen pretest probability with ^{18}FDG PET results.



Swensen: logistic probability model of Swensen et al.[1] ^{18}FDG PET (4 categories): ^{18}FDG PET result in four categories (no uptake, faint, moderate and intense). Swensen + PET: the logistic model combined with ^{18}FDG PET information.

Figure 3. Nomogram using information from the clinical prediction model and ^{18}FDG PET.



The probability of malignancy based on the value of Swensen et al [1] and the ^{18}FDG PET result are indicated in the nomogram. First, the patient's position on each predictor variable scale is defined. Each scale position has corresponding prognostic points located on the "points" scale at the top. These two numbers are then summed to arrive at a "total points" value on the total points axis. A vertical line is then drawn from the total points axis down to the probability to indicate the probability of malignancy.

Figure 4. Bootstrap calibration curve of the clinical prediction model of Swensen et al [1] with the ^{18}FDG PET result (four categories) in 106 patients with SPNs.

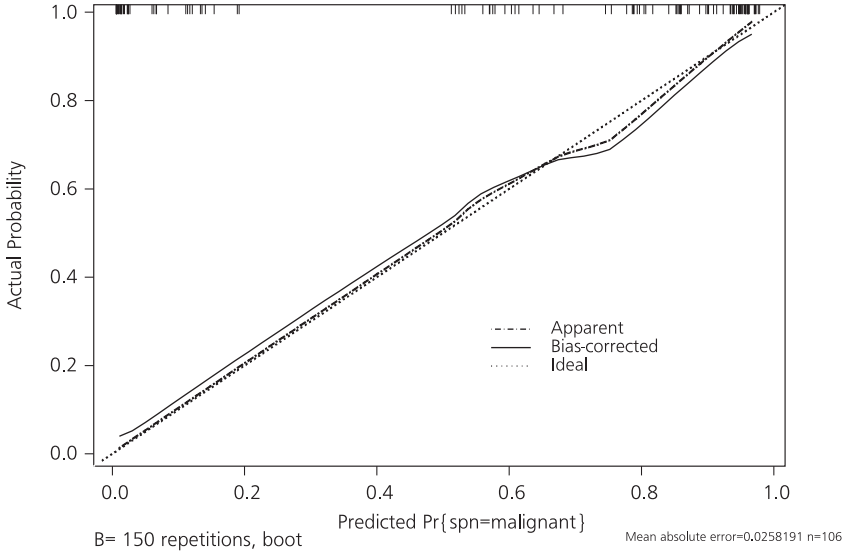


Table 2. Model characteristics

Model	N	DF	AIC[1]	ROC AUC	95% CI	P-value [2]
1. Full model of Swensen et al [2]	419	6		0.83		0.75
2. Validation set for Swensen et al [2]	210	6		0.80		0.62
3. VUMC*: Swensen model	106	1	120.2	0.79	0.70 - 0.87	0.46
4. VU: PET only	106	3	87.0	0.88	0.77 - 0.91	
5. VU: Swensen model+PET	106	4	80.6	0.92	0.87 - 0.97	0.48

* VUMC: present study population. [1] Akaike's Information Criterion. Lower values indicate more desirable models. [2] Goodness-of-fit test p-value [4]

^{18}FDG PET result and the prediction of Swensen et al combined

^{18}FDG PET added to the predicted probability calculated by the Swensen [1] model improves the AUC by 13.6 (95% CI: 6-21; p=0.0003). The fitted function to calculate the probability of malignancy based on the model of Swensen together with ^{18}FDG PET is as follows: probability of malignancy = $1 / (1 + e^{-x})$, with x = -4.739 + 3.691{Probability by Swensen (%)} + 2.322{faint uptake} + 4.617{moderate uptake} + 4.771{intense uptake}. A visual reproduction of the model is given in Figure 3 by means of a nomogram. The corresponding calibration curve displays the relation between the predicted and the actual probability in Figure 4.

Discussion

In 2003, a comprehensive cost-effectiveness decision analysis was published, which included the full spectrum of diagnostic and therapeutic options for SPNs.[2] The first stratification of this analysis was based upon the result of clinical risk assessment as provided by a previously developed multivariate logistic regression model.[1] It was recognized that this model, which was developed in a North-American population with pulmonary nodules discovered between 1984 and 1986 and a prevalence of malignancy of 26.4% [1,7], required additional external validation. The current study provides validation of this clinical prediction model in a sample of patients with radiologically indeterminate nodules collected between 1997 and 2001, with a prevalence of malignancy of 57.5%. Our reported prevalence of 57.5% is more in line with other reports in the literature in which approximately one third of pulmonary nodules were radiologically indeterminate, and, of those, one third of the resected pulmonary nodules were benign.[8,9] However, in spite of differences in prevalence and prevailing local epidemiology of underlying diseases compared to the original data set of Swensen et al [1], the model showed a reasonable fit to our data, indicating that the model is robust. The calibration figure shows that the prediction model tended to underestimate the probability of malignancy, particularly at lower probabilities. Interestingly, in a follow-up study by Swensen et al,[7] in which the probability estimation of four experienced clinicians was compared with

the results of the prediction model, the clinicians tended to overestimate pretest probability particularly at lower values of the predicted probability. Obviously, clinical intuition and judgment of experienced clinicians are most important, but less experienced clinicians may be less accurate and more objective diagnostic techniques are still warranted.

In our study, visual analysis of ^{18}FDG PET proved to be an accurate method of interpretation. The AUC-ROC of ^{18}FDG PET compared to the result obtained with the model of Swensen et al [1] did not significantly (p=0.058) improve predictive value (Table 2). However, the shape of the ROC curve (for the ^{18}FDG PET) especially suggested though that the finding of actual patients having lung cancer (positive predictive value) improved. It could be argued that this was on the border of significance and reflected a type II error. On the other hand, we think that the actual difference of the ^{18}FDG PET-alone model with the prediciton model of Swensen et al [1] is of little clinical importance since the parameters of the model of Swensen are always available prior to ^{18}FDG PET.

Dewan et al [10] found that dichotomized results of ^{18}FDG PET as a single test performed better than the standard criteria developed in a model by Cummings et al [11], including baseline prevalence, size, age and smoking history. However, the series by Dewan et al [10] was smaller, and the comparative model was based on Bayesian analysis combining likelihood ratios of test results that were assumed to be conditionally independent while derived from various sources. Our results suggest that with respect to diagnostic performance, the best results are to be expected from the combined information of clinical assessment and ^{18}FDG PET (*ie,* the AUC-ROC showed a significant improvement (p=0.0003) as did the AIC of the combination model). Limitations of both ^{18}FDG PET studies were their retrospective design as well as the potential of referral bias, which are other reasons for validation of the results.

Clinical prediction rules and modeling can help to set the indication for PET scanning beyond the almost intuitive reasoning that ^{18}FDG PET will be most useful in the 10-50% pretest probability range [2;12]. However, the results of complex decision models obviously depend on several assumptions. For example, it is not clear that the required strict pursuit of histopathological diagnoses can or will be obtained in clinical practice. In fact, it has been claimed that low ^{18}FDG uptake (*ie,* the likely false negative ones) in T1 malignant lesions carries a relatively favourable prognosis,[13] but this has rightly not been accounted for in the model. Whether patients and clinicians will accept the strategy of watchful waiting in such cases remains to be seen.

Even though we are aware that diagnostic accuracy measures are not directly related to patient outcomes, information as provided in the present study will at least help to define whether in individual cases the result of ^{18}FDG PET might affect management. We expect that these limits may not be the same in different clinical situations. Therefore, a logistic model may, apart from calculating posttest probabilities, also help to decide whether ^{18}FDG PET should be performed in an individual patient. After estimating the pretest probability of cancer, the clinician can assess which (if any) ^{18}FDG PET result will push the diagnostic uncertainty beyond required limits. Since our analysis was based upon patients referred for ^{18}FDG PET, we cannot

exclude the possibility of referral bias. Therefore, our model needs validation but since SPN is a major indication for ^{18}FDG PET, this should not be a major problem.

It has been pointed out that ^{18}FDG PET studies in coin lesions contain a variety of criteria by which a PET result can be assigned a positive result. This is of concern when considering implementation of the technique. For practical purposes, the quantitative potential of ^{18}FDG PET is often reduced to semi-quantitative measures like the standardised uptake value (SUV), which basically expresses the concentration of ^{18}FDG uptake in a lesion as a function of the total injected dose. In comparison with visual image analysis, this approach has the conceptual advantage of objectiveness. However, the results of SUV measurements are also prone to heterogeneity due to prevailing differences in data acquisition and reconstruction methodology.[14] A visual analysis of ^{18}FDG PET scan was performed because it has been shown that SUV methodology and implementation is less straightforward than was often assumed. There is still debate about the appropriate normalisations to be used, but, more importantly it has recently been demonstrated that results of SUVs strongly depend on image reconstruction methodology, level of noise, image resolution and region of interest definition, so that its use is highly questionable for generic diagnostic purposes.[15] Moreover one systematic review failed to show that semiquantitative image interpretation improves the accuracy of ^{18}FDG PET.[3] Finally, the lack of standardisation in the current PET literature and practice strongly compromises the theoretical advantage of so-called *objective* measurements. The excellent reproducibility of visual scaling is probably explained by its close association with semiquantitative tumour/nontumour ratios. Our data suggest that visual assessment of ^{18}FDG uptake intensity is a robust method. It is controversial whether attenuation correction improves detection. There is general agreement that localisation of abnormalities can be simplified by this correction, but this is not the issue in coin lesion characterisation. The downside of attenuation correction is a loss of patient throughput by about 30% due to the time needed for the acquisition of transmission scans necessary to obtain an accurate attenuation map. Even though calibration of our data with attenuation corrected scans is required, we do not expect a major impact since our accuracy data nicely fit into the summary ROC curve of the 2001 metaanalysis.[3]

Conclusion

The clinical prediction model of Swensen et al [1] has been proven to have external validity. However, especially in the lower range of its estimates, the model may underestimate the actual probability of malignancy. Visual analysis of ^{18}FDG PET is a robust and accurate method in radiologically indeterminate SPN. The combination of visually read ^{18}FDG PET scans and pretest factors appears to yield the best accuracy. These results can help to adjust diagnostic workup in individual situations.

Acknowledgements

We are very thankful to V Bongers (MD), RAMJ Claessens (MD), MAL Edelbroek (MD), DAKCJM Huysmans (MD), RA Valdés Olmos (MD), DEA Zanin (MD) and A Zwijnenburg (MD) who performed a visual analysis of FDG PET using intensity scales to asses inter-observer correlation.

References

1. Swensen SJ, Silverstein MD, Ilstrup DM et al. The probability of malignancy in solitary pulmonary nodules. Application to small radiologically indeterminate nodules. Arch Intern Med 1997; 157:849-855.
2. Gould MK, Sanders GD, Barnett PG et al. Cost-effectiveness of alternative management strategies for patients with solitary pulmonary nodules. Ann Intern Med 2003; 138:724-735.
3. Gould MK, Maclean CC, Kuschner WG et al. Accuracy of positron emission tomography for diagnosis of pulmonary nodules and mass lesions: a meta-analysis. JAMA 2001; 285:914-924.
4. Hosmer DW, Hosmer T, Le Cessie S et al. A comparison of goodness-of-fit tests for the logistic regression model. Stat Med 1997; 16:965-980.
5. Harrell FE, Jr., Lee KL, Mark DB. Multivariable prognostic models: issues in developing models, evaluating assumptions and adequacy, and measuring and reducing errors. Stat Med 1996; 15:361-387.
6. DeLong ER, DeLong DM, Clarke-Pearson DL. Comparing the areas under two or more correlated receiver operating characteristic curves: a nonparametric approach. Biometrics 1988; 44:837-845.
7. Swensen SJ, Silverstein MD, Edell ES et al. Solitary pulmonary nodules: clinical prediction model versus physicians. - Mayo Clin Proc 1999; 74:319-29.
8. Huston J3, Muhm JR. Solitary pulmonary opacities: plain tomography. Radiology 1987; 163:481-485.
9. Lillington GA. Pulmonary nodules: solitary and multiple. Clin Chest Med 1982; 3:361-367.
10. Dewan NA, Shehan CJ, Reeb SD et al. Likelihood of malignancy in a solitary pulmonary nodule: comparison of Bayesian analysis and results of FDG-PET scan. Chest 1997; 112:416-422.
11. Cummings SR, Lillington GA, Richard RJ. Estimating the probability of malignancy in solitary pulmonary nodules. A Bayesian approach. Am Rev Respir Dis 1986; 134:449-452.
12. Fischer BM, Mortensen J, Hojgaard L. Positron emission tomography in the diagnosis and staging of lung cancer: a systematic, quantitative review. Lancet Oncol 2001; 2:659-666.
13. Marom EM, Sarvis S, Herndon JE et al. T1 lung cancers: sensitivity of diagnosis with fluorodeoxyglucose PET. Radiology 2002; 223:453-459.
14. Boellaard R, Buijs F, Jong H et al. Characterization of a single LSO crystal layer high resolution research tomograph. Phys Med Biol 2003; 48:429-448.
15. Boellaard R, Krak NC, Hoekstra OS et al. Effects of noise, image resolution, and ROI definition on the accuracy of standard uptake values: a simulation study. J Nucl Med 2004; 45:1519-27.

Chapter 5

Prospective use of serial questionnaires to evaluate the therapeutic efficacy of ^{18}FDG PET in (suspected) lung cancer

Gerarda J M Herder
Harm van Tinteren
Emile F I Comans
Otto S Hoekstra
Gerrit J J Teule
Pieter E Postmus
Urvi Joshi
Egbert F Smit

Thorax 2003;58:47-51

Abstract

Background: A study was undertaken to study the effect of ^{18}F-fluorodeoxyglucose (^{18}FDG) positron emission tomography (PET) on diagnosis and management of clinically problematic patients with suspected non-small cell lung cancer (NSCLC).

Methods: A prospective before-after study in a cohort of all 164 patients (university / community settings) referred for ^{18}FDG PET between August 1997 and July 1999. ^{18}FDG PET was restricted to cases where non-invasive tests failed to solve clinical problems. The impact on diagnostic understanding and management was assessed using questionnaires (intended treatment without ^{18}FDG PET, actual treatment choice after ^{18}FDG PET, post hoc clinical assessment).

Results: Diagnostic problems especially pertained to unclear radiological findings (n=112; 63%), mediastinal staging (n=36; 20%) and distant staging issues (n=16; 9%). ^{18}FDG PET findings were validated by reviewing medical records. ^{18}FDG PET had a positive influence on diagnostic understanding in 84%. Improved diagnostic understanding solely based on ^{18}FDG PET was reported in 26%, according to referring physicians, ^{18}FDG PET resulted in beneficial change of therapy in 50%. Cancelled surgery was the most frequent therapy change after ^{18}FDG PET (35%).

Conclusion: ^{18}FDG PET applied as "add-on" technology in patients with these clinical problems appears to be a clinically useful tool, directly improving therapy choice in 25% of patients. The value of increased confidence induced by ^{18}FDG PET scanning requires further evaluation.

Introduction

Medical imaging technology is rapidly expanding and the role of each modality is being redefined constantly. Positron emission tomography (PET) using ^{18}F-fluorodeoxyglucose (^{18}FDG) has emerged as an accurate imaging modality in patients with lung cancer.[1-3] Potential clinical indications include the differential diagnosis of benign versus malignant disease, initial (preoperative) staging, evaluation of suspected recurrences, and follow up after treatment. The use of ^{18}FDG PET in clinical practice is based predominantly on studies of technical performance and diagnostic accuracy.[4;5] To ensure an appropriate use of ^{18}FDG PET, such studies should be followed by an analysis of the impact of ^{18}FDG PET on management decisions, outcomes of care, and cost-effectiveness.

In the northwestern part of the Netherlands where this study was performed, a single PET scanner serves 2.7 million inhabitants, with 50% of its time slots available for clinical purposes. To restrict the use of ^{18}FDG PET to those patients that may benefit most, a program has been developed to evaluate the clinical usefulness of ^{18}FDG PET, investigating the cost-effectiveness of performing ^{18}FDG PET on a routine basis in the preoperative staging of non-small cell lung cancer (NSCLC) [6] and its impact as an "add on" technique in specific problem cases. To measure the clinical value of ^{18}FDG PET in the latter group, we performed a prospective before-after study in a cohort of clinically problematic cases, typically after an extensive conventional work-up. This study design was used during the early studies of computed tomographic (CT) scanning by Wittenberg et al [7] and allows a systematic assessment of the impact of a test on diagnostic understanding as well as on patient management within the clinical context.[8]

Methods

To be eligible for ^{18}FDG PET scanning, patients had to have suspected or proven NSCLC with a diagnostic problem which, according to the referring physician, could not be solved by conventional methods alone and in which the ^{18}FDG PET result might affect patient management. In an attempt to restrict ^{18}FDG PET scanning to such cases, referrals were only accepted after discussion of the case between this physician and the staff nuclear medicine physician in charge at the Clinical PET Centre of the VU University Medical Centre. PET scanning therefore typically followed an extensive conventional work-up. All patients routinely underwent laboratory tests, bronchoscopy, chest radiography and CT scanning extending from the neck to the upper abdomen (including liver and adrenal glands). Additional diagnostic tests were performed in cases with signs and symptoms suggestive of distant metastatic disease. Patients entered in randomised [9] or response monitoring trials [10] were not included in the present report.

Assessment of clinical value

The impact of ^{18}FDG PET on diagnostic understanding, and therapy choice was investigated using three questionnaires (Fig 1). These questionnaires were to be completed by the referring physician before ^{18}FDG PET scanning, shortly after ^{18}FDG PET scanning, and about 6 months after^{18}FDG PET scanning, respectively. In the first questionnaire, information was requested regarding the histological diagnosis (if known), a definition of the current diagnostic problem, a differential diagnostic consideration, the results of diagnostic tests already performed and any planned diagnostic tests. In addition, the referring physician was requested to outline the intended patient management plan if ^{18}FDG PET scanning was not available. The second questionnaire requested information regarding the working diagnosis and planned treatment after ^{18}FDG PET scanning in addition to any diagnostic tests that had been ordered as a direct consequence of the ^{18}FDG PET scan result. In the final questionnaire, the referring physician was requested to convey the final diagnosis and to rate the overall usefulness of ^{18}FDG PET separately in terms of diagnostic understanding and therapy of choice according to the method of Wittenberg et al. [7] This method involves using a 5 point ordinal scale (Box 1), with higher scores representing an increasing positive impact.

All questionnaires were checked for internal consistency between the pre-PET intentional management (questionnaire 1) and post-PET actual management (questionnaire 3). In the case of inconsistencies, the referring physicians were asked to review the cases in question and to revise the overall clinical value rating accordingly and these data were used in the analysis. In the case of PET negative – that is, suspected benign – coin lesions, follow-up was extended beyond 6 months by examining the medical records of these patients.

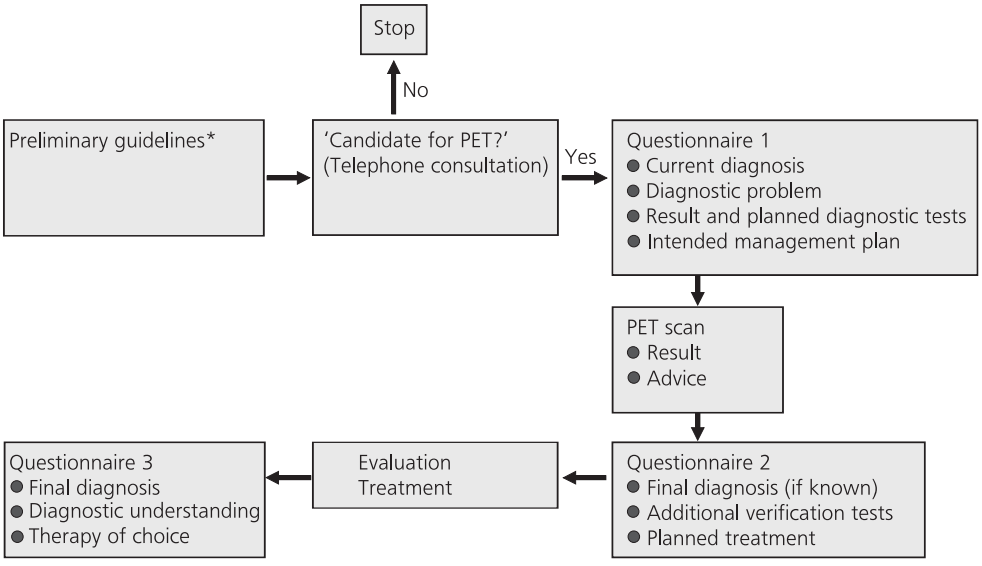


Figure 1. Study protocol. * Suspected NSCLC, diagnostic problem insoluble by conventional imaging, potential impact on patient management.

Box 1. Questionnaire on evaluation of ^{18}FDG PET impact

Diagnostic understanding (DU)

D=1: ^{18}FDG PET confused my understanding of this patient's disease and led to investigations I would not otherwise have done

D=2: ^{18}FDG PET confused my understanding of this patient's disease but did not lead to any additional investigations

D=3: ^{18}FDG PET had little or no effect on my understanding of this patient's disease

D=4: ^{18}FDG PET provided information which substantially improved my understanding of this patient's disease

D=5: My understanding of this patient's disease depended upon diagnostic information provided only by ^{18}FDG PET (unavailable from any other non-surgical procedure)

Treatment choice (TC)

T=1: ^{18}FDG PET led me to choose treatment which in retrospect was not in the best interests of the patient

T=2: ^{18}FDG PET was of no influence in my choice of treatment

T=3: ^{18}FDG PET did not alter my choice of treatment but did increase my confidence in the choice

T=4: ^{18}FDG PET contribute to a change in my chosen treatment but other factors (other imaging tests, other diagnostic tests, changes in patient status) were equally or more important

T=5: ^{18}FDG PET was very important compared with other factors in leading to a beneficial change in treatment

Management changes

Treatment (management) changes were considered "major" if treatment changed from one modality to another – for example, from medical to surgical/radiation/ no treatment or vice versa [11] – and "minor" if treatment changed within a modality – for example, altered medical, surgical or radiotherapy approach.

PET imaging

Whole body ^{18}FDG PET scans were performed with a dedicated PET scanner (ECAT EXACT HR+, CTI/Siemens). Emission scans, typically extending from mid-skull to mid-femur, were acquired in 2D mode, approximately 60 minutes after intravenous injection of 370 MBq (10 mCi) 18FDG. Patients were asked to fast for at least 6 hours prior to the PET study. Oral intake of water was encouraged.

^{18}FDG PET scans were corrected for decay, scatter and randoms. Scans were reconstructed as 128x128 matrices using filtered back projection with a Hanning filter (cut-off 0.5 cycles/pixel) resulting in a transaxial spatial resolution of 7 mm at full width half maximum. If possible, CT scan data were used for more precise anatomical localisation of ^{18}FDG PET abnormalities suspected as being malignant.

Referring physicians were informed by telephone of the result of the ^{18}FDG PET scan and an advice to the next step. Clinicians were urged to verify clinically decisive PET findings by conventional means (histology, imaging, follow-up) and to ignore unconfirmed hot spots. ^{18}FDG PET findings were retrospectively validated by examination of the medical records of

the included patients. Histopathology and clinical follow up findings that showed a benign or malignant course were considered as a valid reference test.

Statistical analysis

Differences in diagnostic understanding or treatment choice between the three indications were tested by means of a two sided Kruskal-Wallis test. Wilcoxon-Mann-Whitney test was used to test differences between two samples. Changes in treatment plans before and after ^{18}FDG PET were tested by the marginal homogeneity test.[12]

Results

During a 23-month inclusion period, 179 patients with suspected NSCLC were referred for ^{18}FDG PET scanning. The referring physicians included pulmonologists (76%), oncologists (7%), internists (6%), radiotherapists (6%), neurologists (3%) and surgeons (1%) from 21 different university and community hospitals. Questionnaires were returned from 178 (99%) patients and a fully completed questionnaires (all questions answered) was obtained for 136 (76%) patients. Specifically, questionnaire 1 was fully completed for 83% of the patients, questionnaire 2 for 92%, and questionnaire 3 for 98%. Indications for ^{18}FDG PET could be subdivided in six groups: unclear radiological abnormality (including solitary pulmonary nodules and lung masses, n=112; 63%), staging of the mediastinum (n=36; 20%), distant staging issues (n=16; 9%), response monitoring (n=5; 2.8%), suspected recurrence (n=5; 2.8%), and unknown primary (n=5; 2.8%). The present report focuses on the first three clinical indications.

In these 164 patients, the clinical work-up before ^{18}FDG PET included laboratory tests, chest radiography, CT scan of the chest (including liver and adrenal glands) and bronchoscopy.[13] In patients with distant staging problems (n=16) the work-up before ^{18}FDG PET consisted of bone scintigraphy and radiographic studies in the three patients with clinical concerns about skeletal metastases; CT evaluation of the abdomen typically preceded referrals with suspect adrenal enlargement or liver lesions in which biopsy was considered not feasible or had been inconclusive. In two patients in which chest CT scan had shown additional and indeterminate pulmonary lesions, bronchoscopic examination had been negative and it was not conidered feasible to take biopsy specimens. In five patients with potentially solitary brain metastases, dissemination tests had included CT scanning (brain, chest, liver and adrenal glands) and bone scintigraphy. In general, the work-up of patients with unclear radiological findings before ^{18}FDG PET scanning conformed to national guidelines.[13]

The diagnostic problems concerning mediastinal staging leading to referral for ^{18}FDG PET (instead of invasive mediastinal staging) included former mediastinoscopy, thoracotomy or radiotherapy, indeterminate invasive staging results, medical inoperability, and “to determine

the most appropriate surgical approach". After careful evaluation we were unable to identify a specific reason for choosing ^{18}FDG PET scanning as opposed to mediastinoscopy to determine mediastinal lymph node involvement in 10 patients.

In 29 out of the 179 patients the initially formulated management plans (to be carried out if ^{18}FDG PET had not been available) were not consistent with the final assessment of the impact of ^{18}FDG PET. For example, the physician's written plan before ^{18}FDG PET was to perform a thoracotomy, and a thoracotomy was indeed performed but treatment choise was rated as 5 (^{18}FDG PET was very important compared with other factors leading to a beneficial change in treatment). Such inconsistent assessments were revised by the referring physicians (specifically with respect to the questionnaire 3), and corrected in 28 cases.

Diagnostic understanding

The impact of ^{18}FDG PET on diagnostic understanding was analysed for each clinical indication (Table 1). Overall, ^{18}FDG PET was solely responsible for improved diagnostic understanding (DU=5) in 26% (95% CI 19 to 33) of the patients and substantially contributed to diagnostic understanding (DU=4) in 58% (95% CI 50 to 65). The effect of the ^{18}FDG PET result on diagnostic understanding was confusing and led to additional tests (DU=1) in 3% (95% CI 1 to 6), and had no or little effect (DU=3) in 9% (95% CI 5 to 15). The impact of ^{18}FDG PET on diagnostic understanding was not significantly different for the three clinical indications (p=0.45). There was no significant difference (p= 0.85) in diagnostic understanding ratings between ^{18}FDG PET scans indicating malignancy where the tumour was finally proven to be malignant (true positives) and scans indicating benign disease where the lesion proved to be benign (true negatives). To evaluate the presence of a potential clinical learning curve of incorporating ^{18}FDG PET scanning results, we compared the diagnostic understanding ratings of "early" patients (the first five patients) referred by a particular physician to the ratings of later patients (the sixth and subsequent patients). The ratings in later patients tended to be significantly higher (p=0.0192).

Table 1. The impact of ^{18}FDG PET on diagnostic understanding (DU) ratings (defined in box 1)

	DU=1	DU=2	DU=3	DU=4	DU=5	Missing	Total
Radiological abnormality	3	6	12	61	29	1	112
Mediastinal staging	1	1	1	21	11	1	36
Distant staging	0	0	2	10	2	2	16
Overall	4	7	15	92	42	4	164

Diagnostic accuracy

Of the patients referred to resolve unclear radiological findings, 76 patients had a positive ^{18}FDG PET scan result which proved to be true positive in 68 patients (89%). Thirty six patients had a negative scan reading –that is, no focally enhanced ^{18}FDG uptake suspicious for malignancy–

which proved to be correct (true negative) in 34 patients (94%) either by "wait and see" policy (n=32) or surgery (n=2). The mean duration of follow up in these patients was 20 months (range 6-36). In two patients the ^{18}FDG PET scans proved to be false negative. These false negative cases included a patient with pulmonary fibrous tumour (the patient underwent a curative pneumonectomy) and a patient with mantle cell lymphoma (diagnosed 1 year after the ^{18}FDG PET scan). In one patient the indeterminate solitary pulmonary nodule proved to be true positive at surgery but ^{18}FDG PET was found to have missed micrometastatic involvement of mediastinal lymph nodes.

Of the patients referred for mediastinal staging, 24 patients had a positive ^{18}FDG PET scan result of which 22 were proven to be true positive as shown by pathology in 16 patients and by follow up in six patients; one was proven to be false positive (as shown by pathology) and one patient was lost to follow up. Eleven patients had negative scan results which were found to be true negative in 10 patients (as shown by pathology in six patients and by follow up in four: mean time from PET to last chest radiograph or CT scan was 15 months, range 13-17). In one patient the PET scan was found to be false negative (as shown by pathology). In one patient the scan trajectory did not include the mediastinum due to claustrophobia.

Of the patients referred because of distant staging issues, 10 were found to be true positive (as shown by pathology in six patients, follow up in two, and radiology in two). Six patients proved to have a true negative ^{18}FDG PET scan as shown by follow up in five patients (mean time of follow up 6 months, range 6-6). In one patient the ^{18}FDG PET result proved to be false negative (bone metastases).

Management changes

In 162 of the 164 cases studied explicit provisional therapeutic plans had been stated before ^{18}FDG PET. In 103 patients this involved surgery. After ^{18}FDG PET, surgery was the treatment most commonly abandoned (Table 2). ^{18}FDG PET contributed to a decision to forego surgical treatment in 36 patients (35%; 95% CI 26 to 45) in whom it had been provisionally planned. Of the patients in whom surgery was not the proposed treatment before ^{18}FDG PET (n=59), seven patients subsequently underwent surgery. In these patients the intended therapy had been observation in

Table 2. Treatment changes after PET (T=4/5, n=78)

Treatment change	No. of patients
Surgery to	
Radiotherapy	6
Chemotherapy	11
Observation	18
Radiotherapy to	
Surgery	1
Chemotherapy	2
Observation	3
Chemotherapy to	
Surgery	2
Radiotherapy	0
Observation	2
Observation to	
Surgery	3
Radiotherapy	4
Chemotherapy	1
Minor changes within	
Surgery	14
Radiotherapy	9
Chemotherapy	2

four patients, chemotherapy in two patients, and radiotherapy in one patient. There was a significant change in terms of the "impact" of treatment for the patient, mainly toward a less aggressive approach (surgery→chemo-/radiotherapy→observation; p=0.0001). The impact of ^{18}FDG PET on treatment was divided into major or minor changes as outlined previously. PET was responsible for changes of choice of treatment that were major in 55 patients (66%; 95% CI 55 to 76) and minor in 28 patients (34%; 95% CI 24 to 45).

Post hoc evaluation of treatment choice

The impact of ^{18}FDG PET on treatment choice was analysed for each scan indication (Table 3). According to the attending physician, ^{18}FDG PET was the most important factor leading to a beneficial change of treatment (TC=5) in 45 of 159 patients (28%; 95% CI 21 to 35) patients and contributed to such change (TC=4) in 34 (21%; 95% CI 15 to 28).

Of the 134 cases in which the physician reported increased diagnostic understanding, therapeutic plans remained unchanged in 59 cases (44%). No significant differences in changes of treatment choice for the three different indications were found (p=0.65). Treatment choice ratings after PET scanning indicating malignancy when the suspected lesion was indeed found to be malignant were not different from scans indicating a benign lesion found to be benign (p=0.27). Like diagnostic understanding, the treatment choice ratings were significantly higher for later patients than for early patients (p=0.037).

Table 3. The impact of ^{18}FDG PET on patient management and its clinical assessment (treatment choice (TC) ratings as defined in box 1)

	TC=1	TC=2	TC=3	TC=4	TC=5	Missing	Total
Radiological abnormality	1	16	42	21	30	2	112
Mediastinal staging		3	11	10	10	2	36
Distant staging		3	4	3	5	1	16
Overall	1	22	57	34	45	5	164

Discussion

A new test that appears to be more accurate than the standard ones will generate a clinical demand, even if it's effect on clinical outcome measures is still unclear. With scarce technology like ^{18}FDG PET overconsumption may result precluding general accessibility. Evidence-based guidelines for routine use are therefore needed, so that the available scanning capacity can be adjusted to the expected demand. However, guidelines aim at the average patient and may not be applicable in specific situations. In this prospective, multicentre before-after study the reported clinical impact of ^{18}FDG PET as an "add-on" technology to solve diagnostic problems in patients with suspected NSCLC was considerable. Clinical compliance with the ^{18}FDG PET results was high, and ^{18}FDG PET was reported to have led to beneficial management changes

(TC≥4) in 50% of the patients in the three clinical situations investigated. In addition, a positive influence on diagnostic understanding (DU≥4) by ^{18}FDG PET was observed in 84% of the patients. Put in a more conservative way, ^{18}FDG PET proved to be the key diagnostic tool in one of every four patients referred for ^{18}FDG PET (DU/TC=5).

Interestingly, we observed an increasing appreciation of ^{18}FDG PET over time. Even though other explanations may also be valid, individual consultation and feedback as done in our setting, is known to improve patient referral patterns.[14]

Interpretation of the classification of "important contribution" to treatment choice by ^{18}FDG PET (TC=4) is not straightforward. It is recognised that, in most clinical situations, decisions are made on the basis of a number of factors. Patient management depends on the preoperative assessment of the probability of disease, which is a joint function of multiple diagnostic indicators such as signs, symptoms and test results together with the effectiveness of the invasive procedures that follow them. This complicates the assessment of the contribution of a single test to a change in patient management. Even though the phrasing of the "contributive" ratings (DU/TC=4) may benefit from accentuation, such positive perceptions may always contain a spectrum of clinical relevance which is difficult to translate into outcome measures. The assessment of the true value of "contributive" rather than directly decisive ^{18}FDG PET findings (TC=4 v TC=5) is therefore best done in a randomised study design.

Some studies have recently addressed the clinical impact of ^{18}FDG PET. The methodologies and patient spectra were variable, but the reported management changes (65-70% [15-17]) are uniformly higher than those observed as a by-product in accuracy studies (10-59% [18,19]). This underlines the fact that management change is multifactorial and does not merely depend on a single test (such as ^{18}FDG PET). Alternatively, "clinical value" studies may have overestimated the true clinical contribution of ^{18}FDG PET. Firstly, the clinical impact of a new technology depends on the quality of the previous clinical work-up; poorly performed conventional staging before ^{18}FDG PET scanning would overestimate its actual value. We therefore made an effort to restrict ^{18}FDG PET referrals to cases in which conventional investigations had indeed been performed and had failed. As we have shown, this was the case in the majority patients. Further, a retrospective analysis of the pre-PET work-up showed adherence to internationally accepted guidelines in the majority of patients. Secondly, whether a specific test contributed significantly is a matter of judgement, and thus subject to disagreement, error and imprecise measurement.[8] This was, indeed, the case in our study; inconsistencies were identified in 18% of the questionnaire responses. To strengthen the evidence of before-after studies, independent reviewing of the data by experts has been suggested. This has been shown to reduce the presumed benefit of a new technology as assessed with this type of study design.[20] However, such findings may also reflect the heterogeneity of daily clinical practice in which patients are actually diagnosed and treated. Thirdly, unconscious bias of the referring clinicians in favour of the new technology may have affected the results. We cannot rule out that this has occurred but the opposite may also be true. Even though the sample was not randomly chosen, we found no such effect in the

medical records of the cases in which a prolonged follow up was needed and the data were derived from a broad spectrum of hospitals.

The questionnaires used do confirm a distinction between the clinical impact of a test on diagnostic understanding, patient management, and (retrospective) clinical assessment of the appropriateness of these changes. The data clearly show that the perceived benefit of ^{18}FDG PET scanning consists of altered patient management but, to an even greater extent, of increased diagnostic understanding or confidence in cases where patient management was not altered. In their present form, the questionnaires do not allow estimation of the amount of clinical uncertainty. In our opinion, studies such as this may serve to estimate the relative merits of ^{18}FDG PET for different indications within a specific clinical context. If ^{18}FDG PET fails to show clinical impact, the presumed indication for ^{18}FDG PET may be removed from the list, whereas promising results warrant further investigation. Our data do not represent consecutive patients presenting with a similar clinical problem, and as such, our results cannot be extrapolated to imply the routine use of ^{18}FDG PET in all patients with suspected NSCLC. Estimation of the cost-benefit of such an application requires a direct comparison between patients subjected to ^{18}FDG PET and conventional work-up. Such a study is currently ongoing in the Netherlands.

In summary, controlled implementation of ^{18}FDG PET as a "last resort" diagnostic modality improved patient management in at least 25% of clinically problematic cases with suspected NSCLC. The combination of preliminary guidelines, intensive feedback, and prospective monitoring may promote the effective use of scarce technology.

Acknowledgement

The authors thank A. Kalwij and C. Karga (secretaries, Clinical PET Centre) for collecting all the questionnaires.

References

1. Dwamena BA, Sonnad SS, Angobaldo JO et al. Metastases from non-small cell lung cancer: mediastinal staging in the 1990s--meta-analytic comparison of PET and CT. Radiology 1999; 213:530-536.
2. Gould MK, Maclean CC, Kuschner WG et al. Accuracy of positron emission tomography for diagnosis of pulmonary nodules and mass lesions: a meta-analysis. JAMA 2001; 285:914-924.
3. Pieterman RM, van Putten JW, Meuzelaar JJ et al. Preoperative staging of non-small-cell lung cancer with positron- emission tomography [comment]. N Engl J Med 2000; 343:254-261.
4. Adams A, Flynn K. Positron Emission Tomography - descriptive analysis of experience with PET in VA. Technology Assessment Program 10, i-A5-4. 1998. Boston, USA.
5. Commonwealth department of health and aged care. Report of the commonwealth review of positron emission tomography. Health access and financing division, Australia, 2000. http://www.health.gov.au/haf/msac
6. Van Tinteren H, Hoekstra OS, Smit EF et al. Effectiveness of positron emission tomography in the preoperative assessment of patients with suspected non-small-cell lung cancer: the PLUS multicentre randomised trial.Lancet 2002;359:1388-93.
7. Wittenberg J, Fineberg HV, Black EB et al. Clinical efficacy of computed body tomography. AJR Am J Roentgenol 1978;13:5-14.
8. Guyatt GH, Tugwell PX, Feeny DH et al. The role of before-after studies of therapeutic impact in the evaluation of diagnostic technologies. J Chronic Dis 1986;39:295-304.
9. Van Tinteren H, Hoekstra O, Smit E et al. on behalf of the IKA-PLUS Study Group. Towards less Futile Surgery in Non Small Cell Lung Cancer? A Randomized Clinical Trial to Evaluate the Cost-effectivensess of Positron Emission Tomography. Controlled Clinical Trials 2001; 22:89-98.
10. Vansteenkiste JF, Stroobants SG, Hoekstra C et al. 18Fluorodeoxyglucose -2-Deoxyglucose Positron Emission Tomography (PET) in the Assessment of Induction chemotherapy (IC) in Stage IIIa0N2 NSCLC: a Multi-Center Prospective Study [abstract]. Journal of Clinical Oncology 2001;20,313a.
11. Seltzer MA, Valk PE, Wong CS et al. Prospective survey of referring physicians to determine the impact of whole body FDG-PET on management of cancer patients [abstract]. J Nucl Med 2000;428.
12. Agresti A. Categorial Data Analysis. Ed. John Wiley & Sons, New York, 1990.
13. van Zandwijk N. [Consensus conference on the diagnosis of lung carcinoma (see comments)]. Ned Tijdschr Geneeskd 1991;135:1915-1919.
14. Eccles M, Steen N, Grimshaw J et al. Effect of audit and feedback, and reminder messages on primary-care radiology referrals: a randomised trial. Lancet 2001;357:1406-1409.
15. Kalff V, Hicks RJ, MacManus P et al. Clinical Impact of (18)F Fluorodeoxyglucose Positron Emission Tomography in Patients With Non-Small-Cell Lung Cancer: A Prospective Study. J Clin Oncol 2001;19:111-118.
16. McCain TW, Dunagan DP, Chin R et al. The usefulness of positron emission tomography in evaluating patients for pulmonary malignancies. Chest 2000;118:1610-1615.
17. Tucker R, Coel M, Ko J et al. Impact of fluorine-18 fluorodeoxyglucose positron emission tomography on patient management: first year's experience in a clinical center. J Clin Oncol 2001;19:2504-2508.
18. Gambhir SS, Czernin J, Schwimmer J et al. A tabulated summary of the FDG PET literature. J Nucl Med 2001;42:4S-8S.
19. Saunders CAB, Dussek JE, O'Doherty MJ et al. Evaluation of fluorine-18-Fluorodeoxyglucose Whole Body Positron Emission Tomography Imaging in the Staging of Lung Cancer. Ann Thorac Surg 1999;67:790-7.
20. Goldman L, Feinstein AR, Batsford WP et al. Ordering patterns and clinical impact of cardiovascular nuclear medicine procedures. Circulation 1980;62:680-687.

Chapter 6

Staging of non-small cell lung cancer and application of ^{18}FDG PET:
A cost modeling approach

Paul Verboom
Gerarda JM Herder
Otto S Hoekstra
Egbert F Smit
Jan HAM van den Bergh
Piet CM van Velthoven
Els WM Grijseels

International Journal of Technology Assessment

Abstract

Background: The presence of (distant) metastases affects the therapy (operation) and prognosis of patients with non-small cell lung cancer (NSCLC). Fifty percent of the operations are futile due to the presence of locally advanced tumour or distant metastases. Therefore, more accurate preoperative staging is required with respect to the outcomes (reduction of futile operations) and costs. This study examines current staging procedure and assesses possible situations for incorporating ^{18}F-fluorodeoxyglucose positron emission tomography (^{18}FDG PET).

Methods: A retrospective analysis was performed to assess actual clinical practice in the staging procedure of 337 patients with NSCLC in two Dutch hospitals. Consequently, by combining these data of actual clinical practice with a literature review, a model was developed to determine the influence of ^{18}FDG PET on the staging outcomes and the costs. In this model the accuracy and costs of PET can be varied as well as the extent of substitution of conventional diagnostic tests by ^{18}FDG PET.

Results: Practice variation was found between the two hospitals with regard to the setting in which the diagnostic staging took place (hospitalisation, outpatient setting) and the extent of the use of mediastinoscopy. This was reflected in the costs and in the number of (futile) operations.

Conclusion: Hospitalisation is the major cost driver in these patients. From a cost viewpoint, the evaluation of ^{18}FDG PET in a strategy after diagnostic imaging but prior to invasive staging seems most optimal.

Introduction

Non-small cell lung cancer (NSCLC) is diagnosed in approximately 9,000 patients per year in the Netherlands. The overall 5-year survival for stage I or II (local disease) is 50 to 60%, while survival decreases substantially in the more advanced stages. In general, the therapy for stage I and II is resection of the tumour; for locally advanced stages, combined systemic and local therapy is optional.[1,2] The strongest prognostic factor for survival is complete resection of the primary tumour. Therefore, it is crucial to optimise staging of the disease in order to initiate optimal treatment and determine resectability. Currently, the staging work-up of patients with NSCLC in the Netherlands encompasses several diagnostic tests. Standard practice includes bronchoscopy to obtain a histological diagnosis of the primary process, followed by imaging tests to select patients with localised disease for surgery, computed tomography (CT) of the chest and abdomen, usually followed by other tests, such as magnetic resonance imaging (MRI) of the brain, ultrasound, bone scintigraphy. According to guidelines, in the patients thereafter held to be candidates for surgery, invasive procedures (mediastinoscopy) are required before resection is planned.[3] Recent studies of NSCLC have shown that positron emission tomography (PET) using ^{18}F-fluorodeoxyglucose (^{18}FDG) is an accurate diagnostic tool e.g. for mediastinal lymph node staging.[1,3-8]
Rational decisions regarding the use of this new technology depend on the assessment of its effects and costs in comparison with the conventional diagnostic staging procedures.[9-11] We have conducted a retrospective study to obtain data on the practice, efficiency and cost of staging cancer in these patients.[3] This article deals with the cost aspects and adds a methodology, in which the value and the costs of ^{18}FDG PET are evaluated, using a modeling approach. Consequently, these findings have been used in the design of a randomised trial.[12]

Methods

Study design

Assessment of current clinical practice encompassed a retrospective analysis of all consecutive patients with (suspicion of) NSCLC diagnosed in 1993 and 1994 in two hospitals: Academic Hospital Vrije Universiteit in Amsterdam (VUMC) and a community hospital, Medical Center Alkmaar (MCA).
Eligible patients underwent one or more of the following staging procedures: bronchoscopy, imaging (e.g. CT scan, bone scan, MRI, ultrasound) and invasive tests (e.g., mediastinoscopy).
Thoracotomy is a diagnostic as well as a therapeutic intervention. It provides information on the presence of malignancy (if this has remained unclear) and on the resectability of the

tumour (diagnosis). If possible, the tumour is radically resected (therapy). In our analysis, we classified thoracotomy as a therapeutic intervention.
Thoracotomy was classified as futile in case of benign lesions, recurrent disease within 1 year after definitive staging and surgery with curative intent, pathologically proven mediastinal lymph node involvement and exploratory ("open and close") procedure for any reason. The number of correctly staged patients and its corresponding costs are the endpoints of the study.

Costs

The costs of the diagnostic strategies were made up of two components: the volume of medical activities, as assessed in our retrospective analysis, multiplied by their monetary valuation. Only direct costs made within the healthcare sector were calculated. Detailed price calculations were made in two hospitals to estimate unit costs reflecting the real use of resources. These calculations included the costs of manpower, materials, equipment and overhead.
Calculation of the cost of the PET scan required additional effort. A fully equipped PET centre as operational in the clinical PET centre of VUMC includes the scanner, the cyclotron and the radiochemicals. The radioactive tracer ^{18}FDG has a physical half life of 110 minutes and is therefore transportable over longer distances; thus PET scanning does not necessarily require a cyclotron in the same hospital (the satellite concept [13]). Typically, 370 MBq of ^{18}FDG is administered to adult patients. The loss of activity due to the time interval between production and administration to the patient needs to be accounted for in the costs. In the present analysis, we excluded the costs of subsequent therapeutic interventions (chemotherapy, radiotherapy).

Results

Patient data

In total, 337 patients were evaluated: 220 patients in the MCA hospital and 117 patients in the VUMC hospital. Patient characteristics (gender, age) were similar in both groups. Final preoperative staging resulted in 45% of all patients with stage I or II at the MCA hospital and 59% at the VUMC hospital.
The total number of futile thoracotomies amounted to 44% in the MCA hospital and 64% in the VUMC hospital. The mean number of (invasive) diagnostic tests between two patients groups (surgery vs. no surgery) in both hospitals is summarised in Table 1. These diagnostic tests included all imaging as well as the invasive tests (bronchoscopy and mediastinoscopy).
No significant differences in the mean number of diagnostic tests were seen between both patient groups and both hospitals. However, the number of patients who underwent invasive diagnostic testing (mediastinoscopy) differed between the various patient groups in both

Table 1: Patient characteristics and hospital practice variation in both hospitals

	MCA (n = 220)	VUMC (n = 117)
Age (mean)	66	63.7
Male gender (%)	177	94
Final pre-operative staging:		
Stage I/II (%)	99 (45)	69 (59)
Stage III/IV (%)	121 (55)	48 (41)
No. of operations in stage I/II patients (%)	85 (86)	59 (86)
No. of futile operations (%)	37 (44)	38 (64)
Operated patients and mean number of diagnostic tests	4.9	5.3
Not operated patients and mean number of diagnostic tests	4.7	5.6
Operated patients and mediastinoscopy (%)	57 (67)	9 (15)
Not operated patients and mediastinoscopy (%)	37 (27)	4 (7)

Table 2: Mean costs prices of the various staging procedures (€) and PET

A: diagnostic tests	Mean cost price (€)*
Hospital day	216
Intensive care day	1,163
Ultrasound abdomen	68
CT scan	123
Bone scan	243
MRI	185
Brochoscopy (flexible)	62
Bronchoscopy (rigid)	167
Mediastinoscopy	361
^{18}FDG PET	1,588

* Mean prices calculated in the two participating hospitals (1 € = 2,2037 DFL)

hospitals. Mediastinoscopy was performed in the majority of patients in the MCA hospital, which reflected the practice variation.

Table 2 shows the full costs of the diagnostic tests and admission days necessary for staging and surgery (intensive care days, hospital days). These prices are the mean prices as calculated in both participating hospitals.

Consequently, we calculated the total costs of the various staging procedures for all patient groups. Table 3 shows the differences in costs as related to the practice variation. For example, in the MCA hospital, preoperative staging takes place in hospital, if there is suspicion of the presence of NSCLC, in order to accelerate the procedure. In the VUMC hospital, however, preoperative staging takes place in an outpatient setting. The difference in the use of a mediastinoscopy between both hospitals is also depicted, but this has less impact on the total cost compared with the hospital length of stay, even if this procedure

Table 3: Cost prices of the various staging procedures in both hospitals (€)

MCA (n=220)	No surgery (n=135)	Unsuccessful surgery (n=33)	Successful surgery (n=52)
Diagnostic staging:			
Number of hospital days	14	15	14.4
Total hospital admission costs	2,604	2,759	2,642
Imaging tests	243	231	214
(non-) invasive test	143	280	267
Total costs (preoperative)	*2,99*	*3,27*	*3,123*
Operation costs			
Number of intensive care days	-	4	4.4
Total costs of intensive care	-	4,803	5,605
Number of hospital days	-	12	12
Total hospital admission costs	-	2,22	2,265
Operation	-	1,477	1,477
Total cost (postoperative)	-	*8,5*	*9,346*
Total costs	**2,99**	**11,77**	**12,474**
VUMC (n=117)	*(n=58)*	*(n=38)*	*(n=21)*
Diagnostic staging:			
Number of hospital days	-	7	6.5
Total hospital admission costs	-	1,671	1,631
Imaging tests	418	318	274
(non-) invasive test	80	109	78
Total costs (preoperative)	*497*	*2,097*	*1,983*
Operation costs			
Number of intensive care days	-	1	2.7
Total costs of intensive care		802	2,854
Number of hospital days		12	13
Total hospital admission costs		2,96	3,226
Operation		1,526	1,526
Total cost (post-operative)		*5,289*	*7,606*
Total costs	**497**	**7,386**	**9,589**

is accompanied by several admission days in the hospital. The total costs for the diagnostic staging procedure is higher in the MCA hospital, due to the hospitalisation and the higher number of mediastinoscopies.

Cost price ^{18}FDG PET

Before introducing PET in the staging process, we calculated the full cost of ^{18}FDG PET. For this calculation, the following information is important: the duration of the ^{18}FDG PET examination, the annual number of ^{18}FDG PET scans, the quantity of radiopharmaceuticals

administered, the staff employed and the mode of obtaining ^{18}FDG (transportation).[13] If we assume six diagnostic PET scans daily, 1,500 scans per year, and a cyclotron in hospital, the average cost of a ^{18}FDG PET scan amounts to €1,588. This reflects the present situation for the Amsterdam PET centre, where the available scan (a full-ring PET scanner) is equally divided between clinical and research applications. It is clear, that if the production and the application change, it will be reflected in the cost of the procedure.

Modeling

From this study, we calculated the difference in average costs per patient in the PET strategies versus the conventional staging process. Decision tree models were constructed for various competing strategies for the inclusion of ^{18}FDG PET in the staging process (Figure 1).

The first strategy (PET 1) included the incorporation of ^{18}FDG PET in the initiation phase of the staging process. In this strategy, all patients are diagnosed by ^{18}FDG PET. The second strategy (PET 2) uses ^{18}FDG PET after medical imaging but prior to invasive staging. The third strategy (PET 3) includes the incorporation of ^{18}FDG PET after medical imaging and mediastinoscopy; patients eligible for surgery are diagnosed by ^{18}FDG PET.

The possibilities for introducing ^{18}FDG PET depend on its accuracy and degree of substitution in the current diagnostic work-up. We must make some valid assumptions beforehand. From a literature review, we know that the accuracy of ^{18}FDG PET is better than in conventional clinical practice. We assume an accuracy of 80% in strategy 1 and 2 and an accuracy of 74% in strategy 3.[14] Furthermore, we assume a constant use of health care resources within each strategy after introduction of ^{18}FDG PET. An additional issue is the question if incorporation of ^{18}FDG PET will result in a substitution or addition of the conventional staging procedures. An advantage of a modeling approach is the possibility to varying the level of substitution of ^{18}FDG PET in the diagnostic work-up process and its consequent impact on the costs. Various levels of substitution are taken into account by varying the percentages, which indicate the levels of substitution, e.g. 0% indicates that all other diagnostic tests (and subsequent hospital admission days) are substituted by ^{18}FDG PET, while 100% indicates that PET is merely added to the staging procedures.

Because the two hospitals differed in staging practice, we applied the calculations to the hospital with the most common applied staging procedure in clinical practice which was the VUMC. Patients with other diagnoses than NSCLC were included in the cost analysis. Patients eligible for surgery, who were not operated due to other medical reasons, were excluded from the accuracy analysis; we therefore included 107 patients. Consequently, by varying the accuracy of PET and the extent of substitution, we were able to determine a cost-neutral strategy.

As is shown in Figure 1, including ^{18}FDG PET leads to a shift in the number of futile operations from 38 to17 and to a corresponding higher number of patients not receiving surgery in the first strategy. ^{18}FDG PET decreases the number of futile operations from 13 to 8 patients in the subgroup receiving a mediastinoscopy, and to 17 in the second and third strategy respectively.

Figure 1. Actual clinical practice of the VUMC and the various ^{18}FDG PET strategies

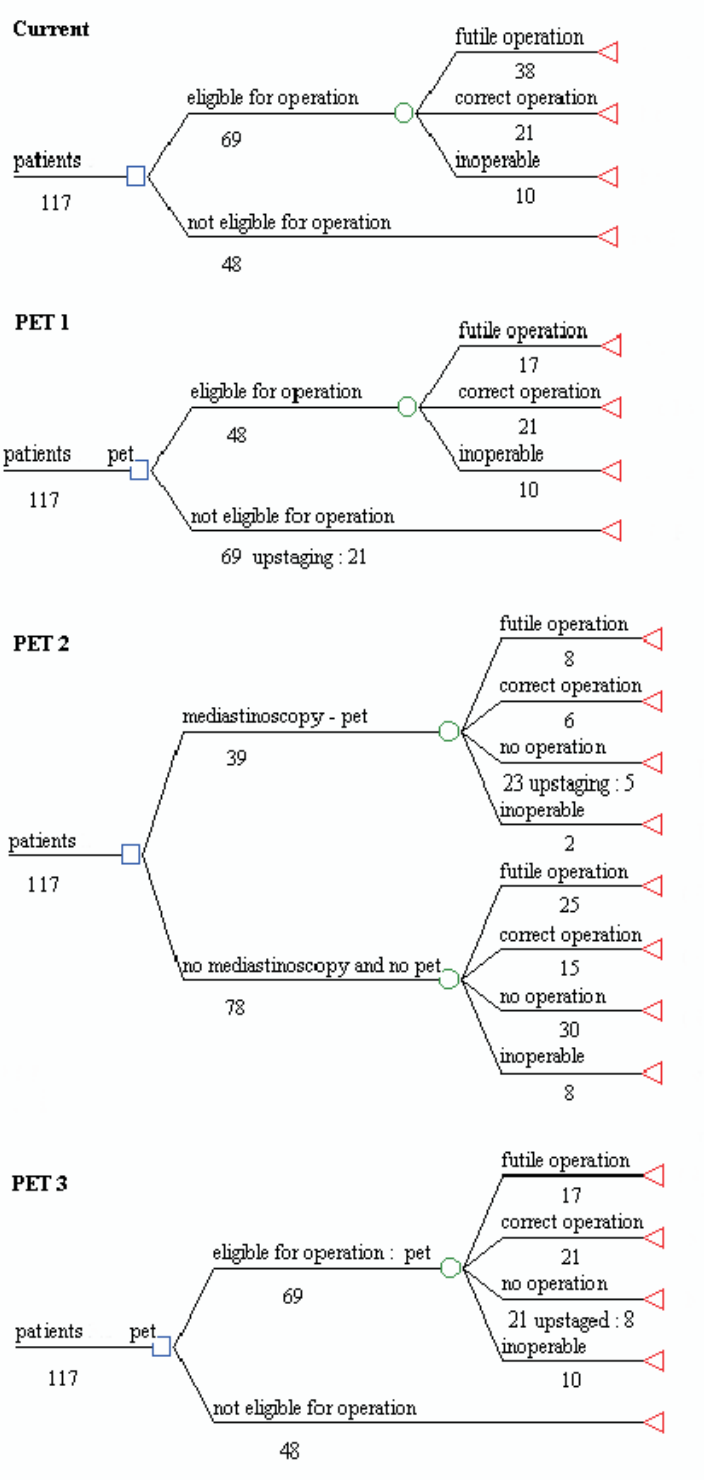
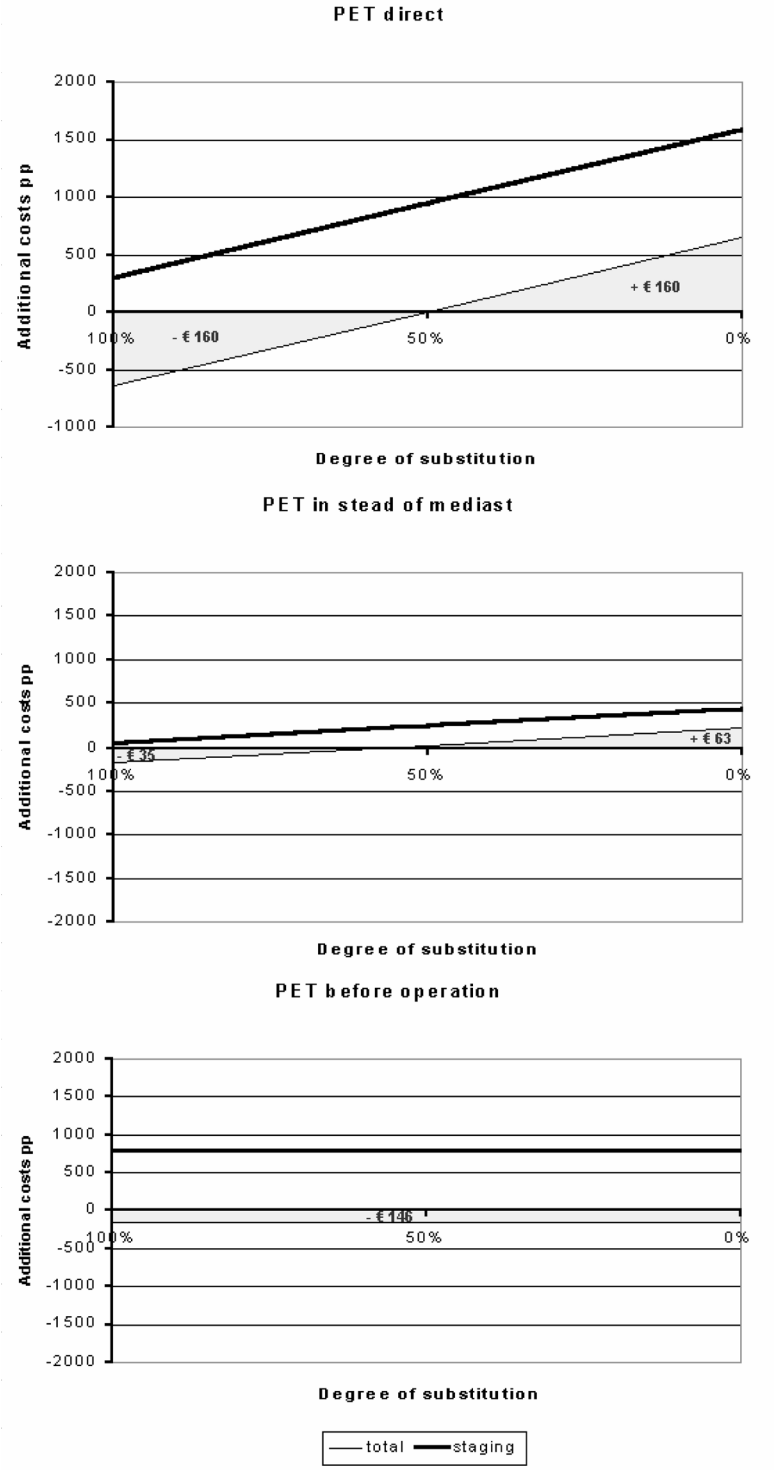


Figure 2. ^{18}FDG PET and percentage substitution (VUMC)



The resulting costs, within a varying degree of substitution, are depicted in Figure 2. On the horizontal axis the degree of substitution is shown, whereas the additional costs are presented on the vertical axis. The bold line shows the additional preoperative staging costs per patient when ^{18}FDG PET is incorporated in the staging process. The thin line shows the total additional costs per patient, including the operation and hospitalisation costs.
Complete substitution (100%) results in an additional € 307 per patient (€ 1,588 PET costs minus conventional staging costs of € 1,281). When PET is completely added (0%), it will result in an additional € 1,588 for each patient.
The downward shifting of the bold to the thin line shows the extent to which PET decreases the number of futile operations. If the accuracy of ^{18}FDG PET equals the conventional staging, both lines coincide. A better accuracy of ^{18}FDG PET results in a further downshifting of the thin line.
As shown in Figure 2, the staging costs are always above the *x*-axis, in contrast to the total additional costs. Its surface under the *x*-axis represents a decrease in costs, whereas the surface above the *x*-axis represents an increase in total costs. The surface of the area under the *x*-axis amounts - € 160, and the surface above the *x*-axis amounts + € 160, resulting in no increase of costs per patient.
With a substitution level of 50%, the additional total costs per patient staged with PET are equal to the costs in the current work-up in the VUMC, as is shown at the intersection of the total cost line with the *x*-axis.
The second strategy has a higher probability resulting in a decrease of costs compared to the first strategy. Complete substitution (100%) results in a decrease of - € 165 per patient. Complete addition (0%) will result in an additional € 221 for each patient.
The surface of the area under the *x*-axis for the total costs amounts - € 35 and the surface above the *x*-axis amounts + € 63, resulting in an additional decrease of € 28 per patient. With a substitution level of 43%, the additional total costs per patient staged with PET are cost-neutral.
For the third strategy, the percentage substitution is not an issue, since all patients have a similar diagnostic work up as in the conventional procedure. Only patients eligible for surgery are diagnosed by ^{18}FDG PET. Therefore, both lines in the figure run horizontal. The bold line is determined by the proportion operations rather than the number of patients eligible for surgery. As shown in the actual clinical practice data, 50.4% of the patients are operated and consequently are eligible for ^{18}FDG PET, which determines the height of the line of € 801. Due to the increased accuracy of ^{18}FDG PET, the total average costs are decreased to € 146.
From our analysis, it is concluded that introduction of ^{18}FDG PET for all patients will lead to a substantial increase in staging costs, which are partially offset by a reduction in futile operations, given a certain amount of substitution of the current diagnostic work-up. ^{18}FDG PET prior to invasive staging will lead to a more limited increase in staging costs, which are offset at a lower level of substitution of other diagnostic tests and therefore seem to be more

efficient. Reserving ^{18}FDG PET only for patients eligible for surgery will always lead to an increase in costs.

Introduction of ^{18}FDG PET will inevitably lead to an increase in staging costs, irrespective of the strategy. By declining the number of futile operations, the total costs can be contained. Certain possibilities and risks accompany each strategy. ^{18}FDG PET in strategy 1 leads to an important increase in costs, but also to an important decrease in total costs. The cost risks in strategy 2 are small, but are accompanied by a small reduction in futile operations. For both strategies, the accuracy and amount of substitution are the major risks. For the third strategy, only the accuracy of ^{18}FDG PET weighs. Finally, the cost price of ^{18}FDG PET in the model can be varied, using a sensitivity analysis approach.

Discussion

PET was developed initially by scientists at Washington University in the late 1970s. PET promised the possibility of imaging physiological changes in the brain, and was seen initially as a unique tool for physiological research. Clinical applications were anticipated. By 1981 the first commercial PET scanner was available. Despite the high price of the technology ($2 to 3 million in 1981), application of the technique in the study of tumours, epilepsy, and cardiology appeared to be promising, especially since the tracer ^{18}FDG proved to have a high affinity for relevant tissues as well as a favourable biodistribution. But the costs and effectiveness of its use in routine clinical practice remains unknown. Recently, the UK standing Group on Health Technology concluded that existing evidence on the diagnostic accuracy of ^{18}FDG PET scan is limited, because the clinical trials were liable to bias.[12] Furthermore, evidence on the cost-effectiveness of PET in various clinical indications is lacking. There is no good insight on how ^{18}FDG PET will effectuate cost-effectiveness in the diagnosis, prognosis and management of patients.

Decisions on the deployment of the ^{18}FDG PET technique are also being made in the Netherlands. Prior to the performance of a randomised clinical ^{18}FDG PET trial, we have formulated a step-to-step course to evaluate the introduction of new diagnostic technologies in general.

The first step encompasses signalling of (in-) efficiency in the existing clinical practice. From our analysis, it was concluded that the percentage of futile operations was very high, that the diagnostic work up varied considerably, concentrated on the number of mediastinoscopies performed and that the hospitalisation days constituted the major cost driver.

However, is it possible to decide whether there is sufficient inefficiency at the current practice to justify the introduction of a new test? If not, one could write a recommendation immediately with this outcome. If there is sufficient inefficiency, we turn to step 2, in which literature analysis on accuracy of the new technology plays a major role. Combining the clinical practice

and the accuracy considerations, this technology can already be recommended for certain indications, as was shown in our analysis. Hereafter, one could perform prospective cost-effectiveness research (step 3), ideally in trial form (randomised). Setting up a randomised controlled trial is difficult and expensive; this should only be contemplated with high chances of health gain and shifting costs, and at specific indications that have the highest chance of occurring the disease. Stage 4 emphasises scenario analysis for testing robustness of assumptions and for extrapolating to other relevant situations (*eg,* additional cost-research with focus on the Dutch situation). The final step includes setting up and implementing guidelines in co-operation with professional associations. This step marks the transition to policy development of pure scientific to more social considerations. This means, for example, that introducing a new diagnostic technology involves other factors, such as clinical relevance, minimisation of discomfort, minimisation of risks, and the time within medical information is available.

Our study can be regarded as an important baseline study for the evaluation of new technologies. By evaluating the actual clinical practice, we noticed a considerable practice variation in the staging process. By analysing the introduction of PET by using a model, we conclude that the greatest impact of PET in effectiveness and costs lies in the second strategy of the staging process.

Acknowledgements:

The authors want to thank, Prof. F.H.H. Rutten and Prof. Raftery for comments on a precious version. Furthermore we want to thank Felicia Barbieri and Peter Blank for helping us calculate the costs in both hospitals.

This study was made possible by a grant of the Dutch ministry of Health for developing cost-effectiveness based guidelines.

References

1. Keller SM, Adak S, Wagner H et al. A Randomized Trial of Postoperative Adjuvant Therapy in Patients with Completely Resected Stage II or IIIa Non-Small-Cell Lung Cancer. New Engl J of Med 2000;343:1217-22.
2. Smit EF, Mourik JC van, Slotman BJ, Postmus PE. De plaats van chemo (radio) therapie voorafgaand aan operatie bij niet-kleincellig longcarcinoom in stadium IIIa (lokaal niet-resectabele tumor). Ned Tijdschr Geneeskd 1999;143:1992-6.
3. Herder GJM,Verboom P, Smit EF et al. Staging of non-small cell lung cancer in two large Dutch hospitals. Eur J Resp Dis. 1999;14;446.
4. Comans EFI, Hoekstra OS, Hoekman K, et al. Meerwaarde van positronemissietomografie met als tracer fluor-18-deoxyglucose (FDG-PET) bij klinische probleemgevallen in de oncologie. Ned Tijdschr Geneeskd 2000;144:1520-8.
5. Dwamena BA, Sonnad SS, Angobaldo JO, et al. Metastases from non-small cell lung cancer: mediastinal staging in the 1990s-meta-analytic comparison of PET and CT. Radiology 1999;213:530-6.

6. Lowe VJ, Fletcher JW, Gobar L. et al Prospective investigation of positron emission tomography in lung nodules. J Clin Oncol 1998;16:1075-84.
7. Mijnhout GS, Pijpers R, Hoekstra OS et al. Primary staging and follow-up of high risk melanoma patients with FDG PET: results of a prospective study of 100 patients. Cancer 1999;85:1199-201.
8. Pieterman RM, Putten JWG van, Meuzelaar JJ, et al. Preoperative staging of non-small cell lung cancer with positron-emission tomography. N Engl J Med 2000;343:254-61.
9. Drummond MF, O'Brien B, Stoddart GL, Torrance GW. Methods for the economic evaluation of health care programmes: second edition. York, UK and Hamilton, Canada: Oxford University Press, 1997.
10. Gambhir SS, Hoh CK, Phelps ME, Madar I, Maddahi J. Decision tree sensitivity analysis for cost-effectiveness of FDG-PET in the staging and management of non-small cell lung carcinoma. J Nucl Med 1996;37:1428-36.
11. Robert G, Milne R. Positron emission tomography: establishing priorities for health technology assessment. Health Technology Assessment 1999;vol.3,No.16.
12. Tinteren H van, Hoekstra OS, Smit EF et al Towards less futile Surgery in non-small cell lung cancer? A randomised clinical trial to evaluate cost-effectiveness of Positron Emission Tomography. Control Clin Trials 2001;22:89-98.
13. Lottes G, Gorschlutter P, Kuwert T, Adam D, Schober O. Costs of F-18-FDG-PET. Experience with a satellite concept. Nuklearmedizin 1998;37:159-64.
14. Dietlein M, Weber K, Gandjour A, et al Cost-effectiveness of FDG-PET for the management of potentially operable non-small cell lung cancer: priority for a PET-based strategy after nodal-negative CT results. Eur J Nucl Med 2000;27:1598-1609.

Chapter 7

Traditional versus up-front ^{18}FDG PET staging of non-small-cell lung cancer: a dutch co-operative randomised study The PET up-front scenario

Gerarda JM Herder
Henk Kramer
Otto S Hoekstra
Egbert F Smit
Jan Pruim
Harm van Tinteren
Emile FI Comans
Paul Verboom
Carin A Uyl - de Groot
Alle Welling
Marinus A Paul
Maarten Boers
Pieter E Postmus
Gerrit JJ Teule
Harry JM Groen
and the POORT study group*

Excepted J Clin Oncol

* The other participants of the POORT study group are listed in the appendix

Abstract

Purpose: We investigated whether application of ^{18}FDG PET immediately after first presentation might simplify staging while maintaining accuracy, as compared to traditional strategy in routine clinical setting.

Methods: At first presentation, patients with a provisional diagnosis of lung cancer without overt dissemination were randomised to traditional work-up (TWU) according to international guidelines or early PET followed by histological/cytological verification of lesions, or imaging and follow-up. Patients with ^{18}FDG avid, non-central tumours without suspicion of mediastinal or distant metastases on PET proceeded directly to thoracotomy. Follow-up in presumed benign lesions was at least 12 months. In patients treated with surgery or neoadjuvant therapy, the quality of staging was measured by comparing the clinical stage to the final stage (combination of peroperative staging and 6 months of follow-up). To investigate test substitution, the number of (non)invasive tests to achieve clinical TNM staging and its associated costs were analyzed.

Results: Between 1999 and 2001, 465 patients (233 TWU, 232 PET) were enrolled by 22 hospitals. The mean (standard deviation) number of procedures to finalize staging was equal in the TWU and PET arms, 7.9 (2.0) versus 7.9 (1.9); p=0.90, respectively. Mediastinoscopies occurred significantly less often in the PET arm. Agreement between clinical and final stage was good in both two arms (Kappa 0.85 vs. 0.78; p=0.07). Costs did not differ significantly.

Conclusion: Up-front ^{18}FDG PET in patients with (suspected) lung cancer does not reduce the overall number of diagnostic test but it maintains quality of TNM staging with the use of less invasive surgery.

Introduction

Evaluation of patients suspected of non-small cell lung cancer (NSCLC) includes diagnosis and staging of the primary lesion and assessment of the extent of locoregional and metastatic spread. Positron emission tomography (PET) using ^{18}F-fluorodeoxyglucose (^{18}FDG) provides useful information in NSCLC staging. The focus of research beyond accuracy measures focused on the added value of PET to conventional work-up if positioned just before to surgery.[1-5] Results of accuracy studies suggest however that application of ^{18}FDG PET up-front in NSCLC diagnostic work-up could also be considered to simplify and improve staging and patient management. ^{18}FDG PET applied early in the diagnostic process might reduce the number of investigations, iatrogenic morbidity and diagnostic delay, and facilitate rapid institution of curative or palliative therapy. Costs of diagnosis and therapy might be reduced if verification of a single decisive lesion suffices to assign appropriate treatment.
The aim of this randomised trial was to investigate whether application of ^{18}FDG PET as an up-front whole body test improves the process of staging patients suspected to have NSCLC without losing accuracy at reasonable costs.

Patients and methods

Patients

Immediately after clinical suspicion of lung cancer had arisen based upon history, physical examination and chest X-ray, patients were invited to participate. Additional inclusion criteria were absence of clinically overt disseminated disease at first presentation, age greater than 18 years and being medically fit for staging and surgery. Exclusion criteria were pregnancy and diabetes. Patients had to give written informed consent according to local ethics committee regulations. Twenty community and two university hospitals recruited patients.

Procedures

Patients were randomly assigned to 'traditional' (TWU) or 'PET up-front' work-up (PET), centrally by computer, by a permuted block design, stratified by institute and (Eastern Co-operative Oncology Group [ECOG]) performance score (0-1 versus 2-3). Patients allocated to TWU underwent imaging (without ^{18}FDG PET) and invasive procedures to establish diagnosis and assess operability and resectability according to international guidelines.[6] Staging in the PET group started with a ^{18}FDG PET scan, which was interpreted in conjunction with available clinical information and chest x-ray. PET scans were performed within one week after randomisation. In either arm suspected locoregional and haematogenous metastases had to be verified by biopsy, or when this was not possible by imaging and follow-up. In the PET arm, invasive confirmation was advised if PET suggested mediastinal lymph node

involvement but no distant metastases. Mediastinoscopy was advised if the primary tumour appeared adjacent to the mediastinum at ^{18}FDG PET, for the reasoning that the spatial resolution of PET does not allow to separate neighbouring mediastinal nodes and primary tumour.[7;8] Surgery was recommended if the primary lung lesion was ^{18}FDG avid without evidence of mediastinal involvement, and distant metastases.[7] Since ^{18}FDG PET is relatively insensitive for brain metastases, clinicians were instructed to perform computed tomography (CT) or magnetic resonance (MRI) of the brain if clinically indicated. A "wait-and-see" policy was acceptable if the primary lung lesion was negative at ^{18}FDG PET. However, when clinicians still wanted further diagnostic information they were instructed to perform staging procedures according to standard practice. All tests and procedures other than ^{18}FDG PET, including treatment and follow-up, were performed in the referring hospitals. Follow-up (up to one year after randomisation) consisted of 3-monthly visits, including at least physical examination, chest x-rays or imaging of indicator lesions in case of "wait-and-see" policy. Figure 1 summarizes the study design. Stages were recorded using the TNM system.[9]

PET imaging and analysis

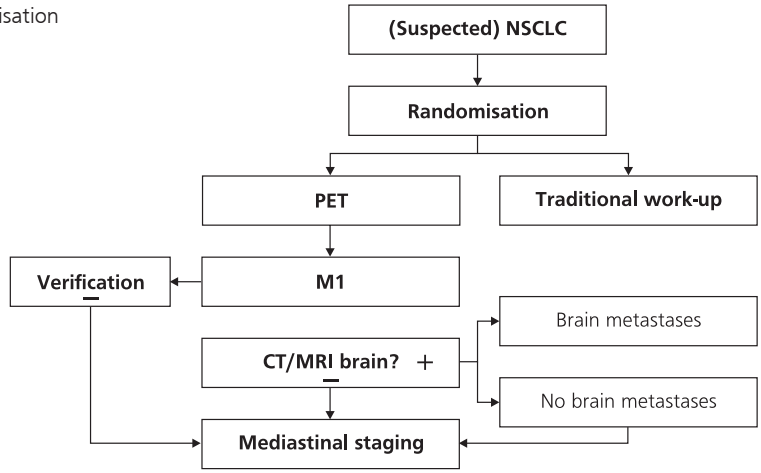
PET scanning was performed in two centres (VU Medical Center, Amsterdam; University Medical Center Groningen) with Siemens ECAT EXACT HR+ scanners (Siemens/CTI, Knoxville, TN, USA). Patients fasted for 6 hours prior to scanning with free access to water. Emission scans were acquired in 2D mode, starting 90 minutes after intravenous injection of approximately 370 MBq ^{18}FDG (if bodyweight > 85 kg: 550MBq). Emissions scans were performed for 5 min/bed-position from knee-joint to skull vertex followed by transmission scanning of the thorax (3 min/bed-position). Scans were corrected for decay, scatter and randoms and reconstructed using ordered subset expectation maximization (OSEM) with 2 iterations and 16 subsets followed by post-smoothing (Hanning 0.5 filter; transaxial spatial resolution 7 mm full-width at half-maximum). In either centre, two experienced nuclear medicine physicians visually interpreted the PET scans. Focally enhanced uptake outside the physiological biodistribution of ^{18}FDG was considered abnormal. Disagreement in interpretations were resolved by consensus, if necessary using a third reader. The final PET report included information on the nature of the primary lesion, the presence of nodal involvement and distant metastases, and concluded with an assessment of TNM stage according to PET and a suggestion for further work-up. Typically, the T classification consisted the likelihood of malignancy of the primary rather than its extension which cannot be assessed reliably at PET.[10] This was communicated to the referring clinicians by phone and confirmed in writing, including a hard copy of PET-images.

Data analysis

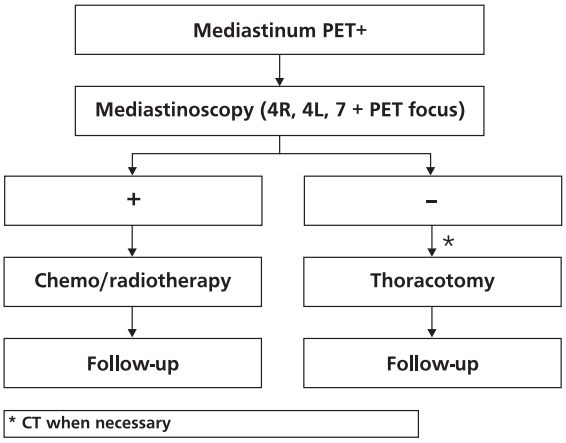
In each patient, clinical stage assigned by the attending clinician was compared to the final TNM stage established at surgery and/or follow-up. The attending physician assigned a 'clinical stage' (c-TNM) using the results of pretherapeutic diagnostic tests. This 'c-TNM'

Figure 1A, B, C. Study design.

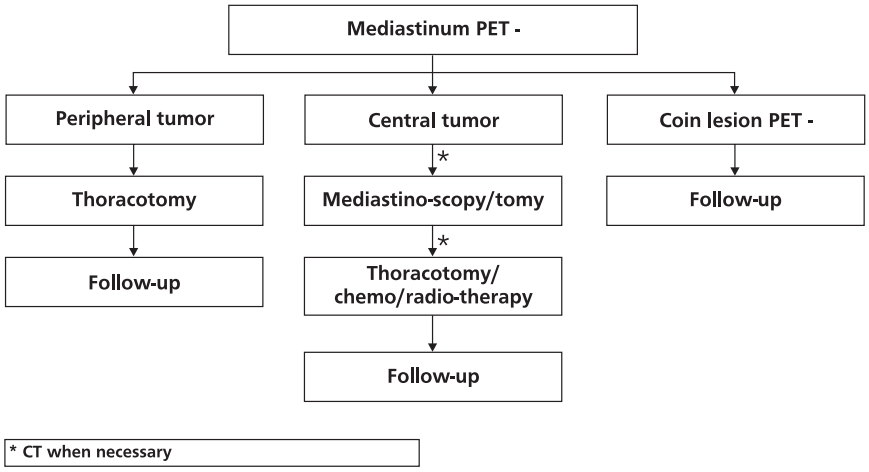
A. Randomisation



B. PET strategy: mediastinum ^{18}FDG positive



C. PET strategy: mediastinum ^{18}FDG negative



stage was compared to the final stage as established by 1) biopsy and/or imaging test results and 6 months follow-up (typically patients with stage IIIB/IV), 2) a combination of surgicopathological staging and 6 months' follow-up (thoracotomy patients), 3) the results at 12 months after randomisation in patients with a provisional diagnosis of a benign primary lesion. In patients who had no thoracotomy due to presumed distant metastasis, and in whom imaging rather than pathology results had been used to establish the final stage, an adjudication committee of three experienced pulmonary physicians reviewed the records to decide whether this clinical classification had been appropriate. We considered staging of lung lesions proving to be metastasis from a primary tumour not identified at clinical work-up as incorrect. If the presenting lung lesion proved to be a single metastasis from a previously known tumour, clinical stage was considered correct provided that no new metastases of the same malignancy became apparent within 6 months. All tests were recorded including procedures to assess resectability and operability (medical fitness). We classified tests as non-invasive (laboratory, functional tests [including ventilation/perfusion scintigraphy, lung and cardiological function tests] and imaging) and invasive (biopsies, surgical procedures).

Outcome measures

Primary outcome measure was the number of tests and procedures to finalize staging and to define operability. Quality of staging was assessed with the number of correctly clinically staged patients compared to the final stage as determined at surgery and/or follow up. Secondary outcome measures included duration of diagnostic processes, morbidity due to complications of diagnostic procedures, and costs of diagnostic and therapeutic processes.

Costs

We calculated total costs from a hospital perspective, implying that only direct medical costs were taken into account. These costs consisted of diagnostic tests (ie, investigations to assess functional operability, staging procedures), therapeutic interventions (ie, thoracotomy, chemotherapy, radiotherapy), outpatient visits and hospital admissions.

The full costs of the various diagnostic and therapeutic procedures were calculated using the micro-costing approach [11], including costs of personnel, materials, depreciation and overheads, calculated as average costs from one general and one university hospital based on 2003 Dutch prices. The costs of ^{18}FDG PET included costs of personnel, depreciation and maintenance, ^{18}FDG, and overheads.[3]

Statistical analysis

Primary endpoint was reduction of number of diagnostic investigations. In an earlier observational study in two participating hospitals, at least 3 (mean, 3.2; standard deviation [SD], 1.6) diagnostic procedures in half of the patients were performed on top of bronchoscopy, chest x-ray, laboratory, lung function and cardiovascular tests and thoracotomy.[12] Here, we considered the PET up-front strategy clinically useful if the proportion of patients needing at

least 3 tests would be reduced from 50 to 30%. Furthermore, we anticipated to include 30% with other histologies (e.g. Small Cell Lung Cancer, benign lung diseases) for which ^{18}FDG PET might have a different impact. Therefore, to sufficiently reliably assess the impact in the patient sample of interest its size was increased by 30%, to a total of 465. Differences in the number of different combinations of tests and duration of diagnostic processes were tested with a *t* test. Decreasing the number of diagnostic procedures should not result in reduction of quality of staging. The latter was measured by overall agreement between clinical and final TNM stage using Kappa statistics. Considering NSCLC patients only, we applied weighted Kappa statistics (Kappa has a maximum of 1.0 with perfect agreement; zero indicates no agreement better than chance). Costs data were compared with two-sided Wilcoxon-Mann-Whitney tests.

Results

Baseline characteristics

Between September 1999 and June 2001, 2114 potential patients were seen in the 22 participating hospitals; of which 465(22%) were enrolled in this study, 233 in the TWU group and 232 in the PET group. One patient allocated to TWU declined further investigations after randomisation. Two patients allocated to PET declined PET and nine allocated to TWU underwent PET. All patients were included in the "intention-to-diagnose" analysis. Baseline characteristics such as age, sex, ECOG performance scores, weight loss, co-morbidity and history of malignancy were well balanced in both groups (Tables 1 and 2). Initial clinical staging did not differ significantly between TWU and PET group (Tables 2 and 3). In the 38 reviewed records of patients with presumed stage IV, in whom imaging rather than pathology results established the final stage, classification was considered to be appropriate.

Table 1. Patient characteristics, demographics

	TWU N=233	PET N=232
Characteristic		
Age (years, mean [SD])	65 (10)	63 (10)
Sex		
Men	155 (67%)	158 (68%)
Women	78 (33%)	74 (32%)
ECOG performance score		
0-1	219 (94%)	227 (98%)
2-3	14 (6%)	5 (2%)
Weight loss >5 %	72 (31%)	68 (30%)
Previous malignancies	33 (14%)	35 (15%)
Vascular comorbidity	68 (29%)	55 (24%)
Final diagnosis of NSCLC	131 (56%)	118 (51%)

Table 2. Distribution of stages of NSCLC at the end of clinical staging period, metastatic and other primary lung malignancies and benign abnormalities.

	TWU N=233*	PET N=232
Benign	27 (12%)	41 (18%)
I/II	85 (36%)	92 (40%)
IIIA	19 (7%)	16 (7%)
IIIB	35 (15%)	24 (10%)
IV	33 (14%)	39 (17%)
Synchronous lung cancer [24]	3 (1%)	3 (1%)
Other primary lung malignancy	26 (11%)	11 (5%)
Metastatic (non-lung)	4 (2%)	6 (3%)

* One patient with missing data. No difference between TWU and PET group, comparing groups by Pearson's Chi-square (nominal) (p=0.24), Kruskal-Wallis (ordered) (p=0.18).

Primary outcome

The proportion of patients requiring at least 3 tests (on top of bronchoscopy, chest x-ray, laboratory measurement, lung functional and cardiovascular tests and thoracotomy [12]) was 52% in the TWU arm compared to 51% in the PET arm (p=0.82). Their total number in order to finalize staging was similar in TWU and PET arm. A mean (SD) of 7.88 (1.95) and 7.90 (1.88) tests (TWU and PET, respectively) was needed for staging in NSCLC (p=0.90). We found no significant difference in the total number of diagnostic procedures.

All patients staged as I/II and IIIa in TWU arm underwent recommended tests (laboratory tests, chest X-ray and CT of chest through liver and adrenals). In patients with stage I/II invasive mediastinal staging was performed in 66% (56/85,95% CI 0.55 to 0.76), and in 45% at least one test procedure (except CT of chest through liver and adrenals) was done to identify distant metastases (38 of 85, 95% CI 0.34 to 0.56). In patients with clinical stage IIIa invasive mediastinal staging was performed in 74% (14/19, 95% CI 0.49 to 0.91), at least

Table 3. Tumor types in patients with other lung malignancies or different metastatic diseases after clinical staging.

	TWU	PET
Other lung malignancies		
SCLC	23	11
Mesothelioma	1	
Sarcoma	1	
Mucoepidermoid carcinoma	1	
Metastatic (non-lung)		
Colorectal carcinoma	1	1
Thyroid carcinoma	1	1
Prostate carcinoma		1
Lymphoma	1	1
Renal cell carcinoma	1	2

one additional test to screen for distant metastases (except CT of chest through liver and adrenals) in 42% (8 of 19, 95% CI 0.20 to 0.67).

Functional tests were evenly distributed among the two groups (Table 4), as were tests aiming at diagnosis and staging. In the PET arm, clinicians used two tests specifically aiming at distant metastases less per ten patients compared to the TWU arm (excluding PET and initial chest CT; mean [SD]: 0.85 [1.09] and 0.63 [1.07] TWU and PET, respectively; p=0.018). The number of patients which required at least one invasive procedure for mediastinal staging was significantly lower (p<0.0001) in favour of the PET arm. The total number of procedures for locoregional staging (bronchoscopy, chest CT, ^{18}FDG PET, mediastinal staging, thoracotomy) was similar in both groups (mean [SD]: 3.8 [1.1] and mean [SD] 3.9 [1.1], for TWU and PET, respectively; p=0.081).

Table 4. Number of tests and procedures for staging lung cancer (n=465)

	TWU	PET	P-value
All tests together (mean (sd))	7.88 (1.95)	7.90 (1.88)	0.90
Functional tests (mean (sd))	2.13 (0.91)	2.23 (0.94)	0.27
All staging tests (mean (sd))	4.75 (1.53)	4.69 (1.52)	0.66
Imaging tests (mean (sd))	3.74 (1.16)	3.80 (1.09)	0.54
Invasive tests (mean (sd))	0.96 (0.95)	0.85 (0.79)	0.18
Invasive tests requiring general anaesthesia (mean (sd))	0.78 (0.85)	0.59 (0.67)	0.0074
≥ 1 invasive test for N staging	92 (39%)	52 (22%)	<0.0001
Thoracotomy	88 (38%)	96 (41%)	0.43

Quality of staging in TWU and PET

Agreement between clinical and final stages was similar (p=0.073) with the two strategies (Kappa 0.85 (95% CI 0.80 to 0.90) in TWU and 0.78 (95% CI 0.72 to 0.84) in PET arm). Adjusted for patients suspected of having NSCLC the weighted Kappa was 0.89 (95% CI 0.82 to 0.95) for TWU and 0.85 (95% CI 0.79 to 0.92) for PET, respectively (Table 5).

Surgery

Of the 233 TWU patients, 79 (34%) underwent 83 mediastinoscopies, and 88 (38%) proceeded to thoracotomy. In the 232 patients in the PET group, mediastinoscopy was performed in 31 (13%) and thoracotomy in 96 (41%) patients. Of patients clinically staged as I/II (TWU n=85, PET n=92), 75 and 78 (TWU; PET) proceeded to surgery. Reasons for not performing surgery in patients classified as I/II were refusal (TWU, n=1; PET, n=2), medical inoperability (6 in either arm), intercurrent disease (TWU, n=1; PET, n=2), death (TWU, n=1), changes in planned preoperative radiotherapy or chemotherapy (TWU, n=1; PET, n=3), unclear diagnostic findings resulting in a wait and see policy (PET, n=1). The number of patients correctly staged as I/II was 65 (76%) of 85 patients in TWU arm and 64 (69%) of 92 patients in PET arm (p=0.4). Staging errors included benign lesions (TWU, n=3; PET, n=2), upstaging

Table 5. Presumptive clinical diagnosis and stage compared with pathological stage or final extent of disease after 6 months follow-up.

TWU	Final stage									
	No data	Benign	I/II	IIIA	IIIB	M1	Metastatic (non-lung)	Synchronous lung cancer	Other lung malignancy	All
	N	N	N	N	N	N	N	N	N	N
Clinical stage										
No data	1	.	.	.	.	.	.	.	.	1
Benign	.	26	.	.	.	1	.	.	.	27
I/II	.	3	65	6	5	3	2#	.	1#	85
IIIA	.	.	1	16	1	1	.	.	.	19
IIIB	.	.	1	.	31	3	.	.	.	35
M1	.	.	.	.	.	33	.	.	.	33
Metastatic (non-lung)	.	.	.	.	.	.	4	.	.	4
Synchronous lung cancer	.	.	.	.	.	.	.	3	.	3
Other lung malignancy	.	.	.	.	.	.	.	.	26	26
All	1	29	67	22	37	41	6	3	27	233

Without previously known tumour

PET	Final stage								
	Benign	I/II	IIIA	IIIB	M1	Metastatic (non-lung)	Synchronous lung cancer	Other lung malignancy	All
	N	N	N	N	N	N	N	N	N
Clinical stage									
Benign	41	.	.	.	.	.	.	.	41
I/II	2	64	11	7	6	.	.	2#	92
IIIA	.	.	15	.	1	.	.	.	16
IIIB	.	1	.	16	6	1#	.	.	24
M1	1	.	.	1	36	1#	.	.	39
Metastatic (non-lung)	.	.	.	.	.	6	.	.	6
Synchronous lung cancer	.	.	.	.	.	.	3	.	3
Other lung malignancy	.	.	.	.	.	.	.	11	11
All	44	65	26	24	50	8	3	13	232

Without previously known tumour

during surgery (TWU, n=10; PET, n=18), metastasis of a non-lung malignancy (TWU, n=2), other primary lung tumour (TWU, n=1, PET, n=2) and upstaging within 6 months follow-up (TWU, n=4; PET, n=6, see the following section).

Follow-up

In either group, one of the primaries presumed benign proved to be malignant during follow-up. Three and six (TWU; PET) patients clinically staged as I/II NSCLC had distant relapses within 6 months after randomisation, (ie, 5%;9 of 177) of all stage I/II patients. Of these, three and four (TWU; PET) relapsed after apparently curative surgery (two did not undergo surgery). Of the patients (TWU, n=19; PET, n=16) with pathologically proven stage IIIA, one and three patients (TWU; PET) were diagnosed as having stage IV disease within 6 months after randomisation.

Secondary outcomes

Staging in the TWU and PET group in 22 different hospitals required a median of 23 days (range, 1 to 193) and 14 days (range, 1 to106) ($p<0.0001$). Patients in the TWU (n=88) and PET (n=96) group underwent thoracotomy after clinical staging at a median of 16 (range, 4 to 116) and 18 days (range, 1 to152), respectively (excluding delayed surgery due to presumed benign lesions in 4 patients). Other reasons for "delayed" surgery included patient refusal (n=1) and co-morbidity (n=1). Morbidity due to staging procedures other than surgery was evenly distributed. Morbidity due to thoracotomy (including cardiac and cerebral events, renal insufficiency, prolonged mechanical ventilation, bleeding, infections) was observed in 41% and 30% (TWU and PET, respectively); p=0.17). Seven TWU (8%) and 11 PET (11%) patients required surgical reintervention for bleeding, broncho-pleural fistula, irradical resection, cardiac tamponade or empyema. Sixteen (18%) and 21 (22%) patients (TWU and PET, respectively) were readmitted to general ward, or intensive care unit (ICU) within 30 days after surgery. Surgical mortality occurred in 4 TWU (4.5%) and two PET patients (2.1%).

Table 6. Mean (sd) cost of diagnostic and therapeutic processes (dollar)

	TWU	PET	P-value
Imaging tests	760 (467)	1964 (369)	<0.01
Other tests including biopsy	904 (1050)	699 (496)	0.01
Surgery	893 (1200)	1018 (1303)	0.28
Hospitalization:			
- pre-operative	1382 (2565)	889 (2406)	0.03
- post-operative	3464 (4986)	3927 (5750)	0.35
- re-admission	246 (1282)	561 (3713)	0.22
Treatment	3701 (5748)	3522 (5776)	0.74
Total	11351 (8479)	12581 (9567)	0.14

Costs

Estimated cost of ^{18}FDG PET was $ 1,557. Costs of diagnostic and therapeutic procedures were $ 11,351 in the TWU group and $ 12,581 in the PET group (p=0.14) (Table 6).

Discussion

^{18}FDG PET has diffused into clinical practice, based predominantly on the basis of diagnostic accuracy studies.[13] To evaluate the effect of early diagnostic use of ^{18}FDG PET on patient management and outcome, to decide whether diagnostic modalities can be replaced by ^{18}FDG PET, randomised clinical trials are required.[2;14;15] In this randomised trial we tested the hypothesis that ^{18}FDG PET up-front strategy reduces the number of tests to classify patients with a high suspicion of lung cancer. The study failed to demonstrate a reduction of the total number of investigations needed for TNM staging. Total costs of staging and therapy were equal in both arms. In patients with stage IV disease a single whole body ^{18}FDG PET including verification of single decisive lesion significantly reduced the number of tests needed for staging.

The randomised design and the participation of physicians and patients from 22 predominantly community based hospitals strengthens the external validity of the study. From data of the National Cancer Registry we estimate that 22% of all patients diagnosed with NSCLC (stage I to IV) in participating hospitals were randomised, which is substantially higher than the 5 to 9% quoted in therapy trials in lung cancer but markedly lower than the 65% in a former ^{18}FDG PET trial.[2;16;17] A limitation of our study was that the level of clinical experience with ^{18}FDG PET was variable among institutions. Further, due to the up-front positioning of ^{18}FDG PET, ^{18}FDG PET scans were not read in conjunction with CT which is known to improve the accuracy of either test.[10] This practice was enforced by the multicentric nature of the study with ^{18}FDG PET and CT being performed 'on-site' (allowing co-reading of scans) in only two of 22 hospitals.

Even though the present study did not aim to measure impact on patient outcomes, it appears that there was a trend towards less futile surgery (benign lesions, per- or postoperative upstaging) in the TWU arm (20%) than in our previous experience.[2;12] In the PLUS trial, where addition of ^{18}FDG PET to TWU was studied, futile surgery was observed in 30% of patients in the conventional arm at 6 months after randomisation. Our ^{18}FDG PET data are difficult to compare with other studies [2; 5; 7;18] since our key issue was substitution rather than added value of ^{18}FDG PET.

In conclusion, even though 'up-front' ^{18}FDG PET in patients with (suspected) lung cancer does not simplify staging, it still provides good quality TNM staging with the use of less invasive surgery. Further research should determine whether up-front positioning of PET-CT (rather than PET and CT alone) might be a cost-effective alternative for current practices.

Reference List

1. Gonzalez-Stawinski GV, Lemaire A, Merchant F et al. A comparative analysis of positron emission tomography and mediastinoscopy in staging non-small cell lung cancer. J Thorac Cardiovasc Surg 2003;126:1700-3
2. Van Tinteren H., Hoekstra OS, Smit EF et al. Effectiveness of positron emission tomography in the preoperative assessment of patients with suspected non-small-cell lung cancer: the PLUS multicentre randomised trial. Lancet 2002;359:1388-1393
3. Verboom P, Van Tinteren H, Hoekstra OS et al. Cost-effectiveness of FDG-PET in staging non-small cell lung cancer: the PLUS study. Eur J Nucl Med Mol Imaging 2003;30:1444-9
4. Verboom P, Herder GJ, Hoekstra OS et al. Staging of non-small-cell lung cancer and application of FDG-PET. A cost modeling approach. Int.J Technol Assess Health Care 2002;18:576-585
5. Viney RC, Boyer MJ, King MT et al. Randomized controlled trial of the role of positron emission tomography in the management of stage I and II non-small-cell lung cancer. J Clin Oncol 2004;22:2357-2362
6. Pretreatment evaluation of non-small-cell lung cancer. The American Thoracic Society and The European Respiratory Society. Am J Respir Crit Care Med 1997;156:320-332
7. Pieterman RM, van Putten JW, Meuzelaar JJ, Mooyaart EL, Vaalburg W, Koeter GH et al. Preoperative staging of non-small-cell lung cancer with positron- emission tomography. N Engl J Med. 2000;343:254-261
8. Vansteenkiste JF, Stroobants SG, De Leyn PR et al. Mediastinal lymph node staging with FDG-PET scan in patients with potentially operable non-small cell lung cancer: a prospective analysis of 50 cases. Leuven Lung Cancer Group. Chest 1997;112:1480-1486
9. International Union against Cancer. TNM Classification of Malignant Tumours, 5th edition ed. New York; 1997.
10. Lardinois D, Weder W, Hany TF et al. Staging of non-small-cell lung cancer with integrated positron-emission tomography and computed tomography. N Engl J Med 2003;348:2500-2507
11. Gold MR, Siegel JE, Russell LB et al. Cost-effectiveness in health and medicine New York; 1996.
12. Herder GJ, Verboom P, Smit EF et al. Practice, efficacy and cost of staging suspected non-small cell lung cancer: a retrospective study in two Dutch hospitals. Thorax 2002;57:11-14
13. Adams A and Flynn K. Positron Emission Tomography - descriptive analysis of experience with PET in VA. Technology Assessment Program 10, i-A5-4. 1998. Boston, USA.
14. Guyatt GH, Tugwell PX, Feeny DH, Haynes RB, Drummond M. - A framework for clinical evaluation of diagnostic technologies. CMAJ 1986;134:587-594
15. Van Tinteren H, Hoekstra OS, Boers M. Do we need randomised trials to evaluate diagnostic procedures? Eur J Nucl Med Mol Imaging 2004;31:129-31
16. Spiro SG, Gower NH, Evans MT et al. Recruitment of patients with lung cancer into a randomised clinical trial: experience at two centres. On behalf of the Big Lung Trial Steering Committee. Thorax 2000;55:463-5
17. Stephens R, Gibson D. The impact of clinical trials on the treatmentof lung cancer. Clin Oncol 1993;5:211-9
18. Reed CE, Harpole DH, Posther KE et al. Results of the American College of Surgeons Oncology Group Z0050 trial: the utility of positron emission tomography in staging potentially operable non-small cell lung cancer. J Thorac Cardiovasc Surg 2003;126:1943-51

Research support

This study was supported by a grant of ZonMw Program The Netherlands Organization for Health Research and development, Health Care Efficiency Research Program.

Appendix

The POORT study group participants
J Berkovits
B Biesma
FJM van Breukelen
PC Dekker
P Eppinga
AHM van der Heijden
PJH Janssen
C Jie
JP Kaajan
W de Kanter-Koppenol
J Kraan
HB Kwa
EM Mathijssen-de Ruiter
J Nabers
GD Nossent
J Schouwink
RALM Stallaert
WFM Strankinga
PJ Teengs
MA Sulzer
G Visschers
GM Wittebrood-Smit
N van Zanwijk

Summary and discussion

Samenvatting en discussie

Chapter 1 comprises an introduction and outline of this thesis. In the Netherlands lung cancer is still one of the most common cancer in males. On the basis of the 2000 figures of the Netherlands, the incidence of lung cancer was 79.8 in men and 27 in women per 100.000 per year. Survival depends on stage, and survival at 5 years for all stages combined in the Netherlands is 10%. The majority of lung cancer patients have tumours histologically classified as NSCLC (84%). Therapy of NSCLC depends on stage, and staging depends on the quality of diagnostic techniques. A new diagnostic modality like ^{18}F-fluorodeoxyglucose positron emission tomography (^{18}FDG PET) in non-small cell lung cancer (NSCLC) was introduced and evaluated in accuracy and impact on patient management and outcome.

The work presented here is an integral part of the research on evaluation of the role of ^{18}FDG PET in NSCLC as conducted in the VU university medical centre. Typically, this research was done in close collaboration with other national and international university and community hospitals. The scientific approach in this thesis largely follows the framework to study the effectiveness of diagnostic tests described by van Tinteren et al. [van Tinteren, submitted]. This framework starts with an in-depth analysis of the size and nature of the potential residual inefficiency of current daily clinical practice. Then, if literature analysis reveals deficiencies in accuracy data of the new test, these need to be generated. If merely accuracy data of the new test are available, the next step is to estimate the potential impact of adding the new test to the diagnostic algorithm (the 'clinical value') in a real-life setting. At the same time, modelling studies are done to explore potential cost-effective applications of the new test which can subsequently be put to the test, preferably in a randomised design. Together, these steps should culminate in the development and implementation of an evidence-based guideline. After monitoring of its impact in terms of relevant patient outcomes at a macro-level and assessment of the pertaining residual inefficiency, a new cycle of evaluation might be started which can also account for new trends in management diagnosis.

Applied to the context of staging patients suspected of potentially resectable NSCLC, we conducted the following studies:
Chapter 2 describes the first step in the framework in which we explored the clinical practice, yield and costs of preoperative staging in patients with (suspected) NSCLC in an academic and general hospital during a time period of two years (1993/1994). We found that in nearly 50% of operated patients with NSCLC, surgical treatment failed because of an irresectable tumour or a benign lesion during surgery, recurrence or metastases within 1 year after surgery. During surgery 33 (23% of patients who underwent surgery) were irresectable, 19 (13%) had a benign lesion. Surgery was considered as futile in 18 patients (13%) who developed metastases or local recurrence (n=1, 0.7%) within 12 month following radical surgery. In many patients, the effort of staging proved to be considerable in terms of diagnostic load (mean of 5 diagnostic tests (standard deviation ± 1.5) conducted over a median of 20 days and in 13% more than 6 weeks. In many patients, the effort of staging proved to be considerable in

terms of diagnostic load, duration and cost, and with limited success to prevent patients to be subjected to surgical treatment from which they derived no obvious benefit. Failures relate to the quality of diagnostic work-up at every aspect of the TNM staging system.

As pointed out, one out of 7 unnecessary operations concerned patients who proved to have no malignancy at all. Among array of papers on the accuracy of ^{18}FDG PET in radiologically indeterminate pulmonary lesions summarised in systematic reviews, we identified a lack of data on the accuracy of PET as a function of size. As a first approach, we reported on the performance characteristics of ^{18}FDG PET as a function of pre-test risk assessment in radiologically indeterminate solitary pulmonary nodules (SPN) ≤10 mm on spiral computer tomografy in *Chapter 3*. In a retrospective study, we identified 35 patients with 36 SPN ≤ 10 mm (range 3-10 mm) in diameter at clinical presentation. PET imaging correctly identified 30 of 36 small lesions. One lesion proved to be false negative at PET and in 5 lesions, PET scans proved to be false positive. Specificity was 77% (17/22; 95%CI:0.55-0.92), sensitivity 93% (13/14; 95%CI:0.66-1.0), positive predictive value 72% (13/18; 95%CI:0.46-0.90) and negative predictive value 94% (17/18; 95%CI:0.73-1.0). These data suggests that ^{18}FDG PET imaging could be a useful tool in differentiating benign from malignant SPNs ≤10 mm (range 3-10 mm) in diameter at clinical presentation.

The accumulated PET literature provided ample data on the diagnostic accuracy of PET in coin lesions but surprisingly few on the added value of PET vs. clinical probability estimates, even though such knowledge is a prerequisite for beneficial change in diagnostic understanding and therapeutic impact. Even though in daily practice most clinicians will use an intuitive estimate of likelihood of cancer in individual patients, the interrater variability is unknown and a more objective scoring system would be preferable. Swensen et al. have proposed and internally validated a mathematical formula which expresses the probability of malignancy as a function of 3 clinical (age, smoking, history of cancer) and 3 radiographic (diameter, spiculation, location) variables. In *Chapter 4*, using clinical, CT and PET data obtained in 106 patients referred for PET for this problem, we obtained an external validation of the prediction model. At the same time, we quantified the potential added value of ^{18}FDG PET as a function of its operating characteristics in relation to this prediction model. ROC analysis suggested that the best results are to be expected from the combined information of clinical assessment and ^{18}FDG PET.

To explore and improve our understanding of the potential clinical value of ^{18}FDG PET in coin lesions and in a broader setting of preoperative diagnostic problems, we designed a "before-after study" (*Chapter 5*). Such clinical value studies give information to diagnostic understanding and influences on therapeutic decision making. Evaluation was performed by prospectively using questionnaires just before, immediately after and several months after the learning results of the ^{18}FDG PET scan. Patients were referred to the ^{18}FDG PET centre because

of suspected NSCLC, diagnostic dilemmas such as unclear radiological findings. Increased diagnostic understanding was reported in 84% and management changed appropriate in 59% of cases. Cancelled surgery was the most frequently reported change in management (35%). Of the patients referred to resolve unclear radiological findings, improved diagnostic understanding or beneficial management change was reported in 46 and 71% of cases.

In *Chapter 6*, using the available data on accuracy and costs of diagnostic work-up and therapy, we used a decision modelling approach to assess whether and how ^{18}FDG PET might be cost-effective for routine use in the preoperative staging of patients with NSCLC. Practice variation was found between the two hospitals with regard to the setting in which the diagnostic staging took place and the extent of the use of mediastinoscopy. This was reflected in the costs and in the number of operations. Hospitalisation was the major cost driver in these patients. Introduction of ^{18}FDG PET for all patients will lead to a substantial increase in staging costs, which are partially offset by a reduction in futile operations, give a certain amount of substitution of the current diagnostic work-up. ^{18}FDG PET prior to invasive staging will lead to a more limited increase in staging costs, which are offset by a lower level of substitution of other diagnostic tests and therefore seem to be more efficient. Reserving ^{18}FDG PET only for patients eligible for surgery will always lead to and increase in costs. By declining the number of futile operations, the total costs can be contained. From a cost viewpoint, the evaluation of PET in a strategy after preceding diagnostic imaging but prior to invasive staging seemed most optimal.

Since randomised controlled trials provide most reliable and convincing evidence in estimating the role of ^{18}FDG PET in the staging of NSCLC [1], we performed an RCT [2] investigating the added value of ^{18}FDG PET if implemented just prior to mediastinoscopy/thoracotomy and a second trial (described in *Chapter 7*) to investigate substitutional performance of ^{18}FDG PET if applied completely upfront in the diagnostic algorithm. The research question was whether application of ^{18}FDG PET immediately after first presentation might simplify staging while maintaining accuracy, as compared to the traditional strategy in routine clinical setting. At first presentation patients with a provisional diagnosis of lung cancer without overtly dissemination were randomised to traditional work-up (TWU) according to international guidelines or early ^{18}FDG PET followed by histological or cytological verification of lesions or imaging and follow-up. Between 1999 and 2001, 465 patients (233 TWU, 232 PET) were enrolled by 22 hospitals. Mean (sd) number of procedures to finalise staging was equal in the TWU and PET arms: 7.9 (sd:2.0) vs. 7.9 (sd:1.9); p=0.90. Agreement between clinical and final stage was good in both two arms (Kappa 0.85 vs. 0.78; p=0.07). The costs did not differ significantly. In conclusion the application of ^{18}FDG PET up-front in staging of patients with (suspected) lung cancer carries similar overall quality (of accuracy) as compared to TWU, but does not simplify staging.

Guideline development

In 2001, a regional multidisciplinary committee developed a preliminary guideline on the use of ^{18}FDG PET in NSCLC.[3] Together with the growing availability of ^{18}FDG PET in our region, this resulted in better access to ^{18}FDG PET for this indication, and data from the regional comprehensive cancer centre registration subsequently showed an persistent decrease in the number of lung resections in the order of magnitude as predicted by the first randomised trial.[4]
In 2004, a national multidisciplinary committee developed evidence-based guidelines for diagnostic procedures and management in NSCLC.[5] ^{18}FDG PET is recommended in patients with stage I/II/III disease eligible for curative surgery, after conventional work-up and before invasive mediastinal staging. In case of suspected locoregional or haematogeneous metastases verification is advised. Mediastinoscopy or endobronchial ultrasound (EBUS) or transoesophageal ultrasound (EUS) is advised if the primary tumour appears to be adjacent to the mediastinum at PET, reasoning that the spatial resolution of PET does not allow to separate neighbouring metastatic mediastinal nodes from the primary tumour.[6;7] In case of apparent primary pulmonary ^{18}FDG avid lesions without evidence of mediastinal lymph-node involvement, distant or extrathoracic metastases, surgery is proposed. The implementation of the (new diagnostic intervention or) guideline is currently in progress in the Netherlands. An inventarisation for residual inefficiencies should be performed after adequate implementation of the guidelines and will be performed in future.

The aforementioned guidelines did not address the issue of the potential use of PET in clinicoradiologically indeterminate coin lesions. Therefore, using the available literature we suggest the following approach.

The use of ^{18}FDG PET in clinicoradiologically indeterminate coin lesions $\geq$ 1 cm.

Data of lung cancer screening showed that helical CT is more sensitive than chest radiographs for detecting small nodules.[8-10] Due to these screening sessions and ongoing development of helical CT, more and smaller lung nodules will be detected. Considering the importance of identifying malignant nodules as early as possible (5-year survival in patients with resected non-small cell lung cancer (NSCLC) stage IA that has been resected can be 80% [11;12]), next step in the diagnostic algorithm should be clear. Different strategies can be followed like transthoracic needle biopsy, operation, a wait-and-see policy or ^{18}FDG PET. A careful consideration should be made based on test performance, morbidity and costs. The chance of a non-diagnostic result in transthoracic needle biopsy is significant in lesions smaller than 2 cm, and the risk of a pneumothorax is substantial (3.1- 41.7% [13]). As mentioned in this thesis, 15% of all operated patients showed to have a benign lesion.[14] Since a wait-and-see

policy carries potentially adverse effects on outcome, there is a demand for accurate non-invasive tests to prevent surgical interventions for benign lesions.

Adding ^{18}FDG PET to the diagnostic work-up improves selection of surgical candidates. Pooled sensitivity and specificity of ^{18}FDG PET in solitary pulmonary lesions ≥ 1 cm was 97% and 78%.[15;16] This implicates a probability of malignancy of 1% if pre-test probability is 20% and ^{18}FDG PET scanning is negative vs. a post-test probability of malignancy of 86% if pre-test probability is 80%. The application of ^{18}FDG PET in these patients is related to the chance of changing treatment plans. Choice of treatment depends on probability of malignancy. This depends on clinical and radiological findings. Fischer showed that ^{18}FDG PET should first be applied to populations with a pre-test probability of cancer between 10-50%.[15] A negative ^{18}FDG PET scan with a prevalence of 50% or lower more or less rules out malignant disease.
Applying ^{18}FDG PET in a population with a higher prevalence (more than 50%), a negative PET would give higher post-test probability and further examinations should be performed.
Differential diagnosis of ^{18}FDG PET negative coin lesions include bronchioloalveolar carcinoma, pulmonary carcinoid tumours, metastastases from Grawitz, neuro-endocrine tumours etc.[17;18] In case of a wait-and-see policy, repeat radiological tests should be performed during 18 month. Operation should be performed in case of growth.

With subcentimeter coin lesions, typically being picked up more often nowadays due to low-dose CT lung cancer screening [8;19;20], the potential role of ^{18}FDG PET needs to be investigated further. With respect to accuracy, this can be done relatively simple by performing ^{18}FDG PET scans in selected subsets of patients with unclear CT abnormalities. ROC analysis is required to refine the performance characteristics of ^{18}FDG PET as well as its optimal operating characteristics. The ROC curve in the meta-analysis of Gould et al. revealed a Q point (point on the curve closest to the left upper corner of the graph) corresponding with a sensitivity of 94% and a specificity of 83% (Fig 1).

In a setting with very low pretest probabilities, it can be predicted that applying the standard criteria for ^{18}FDG PET positivity would not be clinically productive (Fig 2). However, these predictions are indeed hypothetical since the ROC curve might be different for small lesions due to partial volume effects. With the development of gating technology in ^{18}FDG PET and ^{18}FDG PET-CT scanners, which might also improve the quantitation of ^{18}FDG uptake in small lesions, it would be desirable to conduct an accuracy study in this context. At the same time, clinically acceptable levels of uncertainty should be investigated, i.e. in patients as well as clinicians.

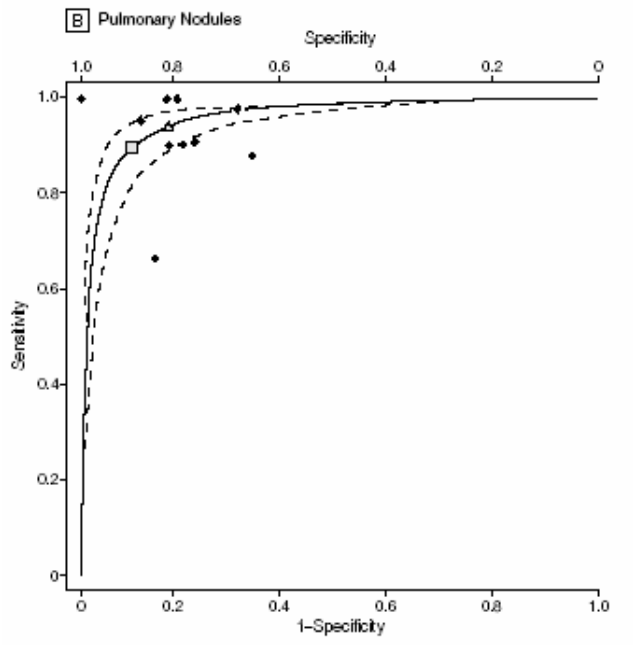


Figure 1. Black diamonds indicate individual study estimates of sensitivity and 1- specificity. Gray squares indicate maximum joint sensitivity and specificity (a global measure of test accuracy) and gray triangles represent the points on the receiver operating characteristic (ROC) curve at which positron emission tomography with ^{18}FDG PET approximately operates in current practice for detecting malignancy in pulmonary nodules. Figure from Gould et al. Jama 2001.[16]

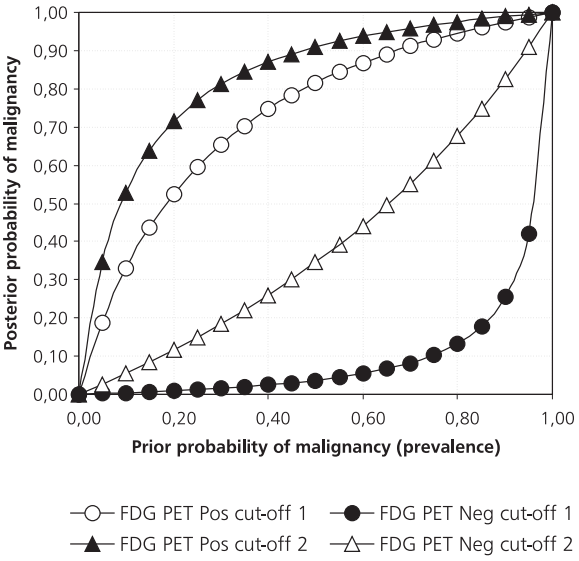


Figure 2. Predicted post-test probabilities of ^{18}FDG PET at different operating characteristics (at cut-off 1 [the standard approach of sensitive PET reading]: sensitivity 0.94, specificity 0.77; at cut-off 2 [more conservative interpretation]: sensitivity 0.50, specificity 0.95).

Future aspects

PET-CT

A new advance in imaging technology is hybrid imaging instrumentation, combining 2 or more technologies such as PET and CT. ^{18}FDG PET and CT provide complementary molecular and anatomic information, respectively, with ^{18}FDG PET adding specificity to anatomic findings and CT offering precise localization of metabolic activity and, in the lung, identification of lesions below the detection limit of PET. Sofar, the acquisition and interpretation of the 2 image sets were done separately. Recently, integrated ^{18}FDG PET-CT systems have become available; these systems provide ^{18}FDG PET and CT images that are acquired nearly simultaneously and are capable of producing superimposed, co-registered images, facilitating interpretation.

The most obvious advantage of PET-CT vs either technique alone in NSCLC is to distinguish tumour from atelectasis.[21-23] Even though there have been claims that PET-CT might add to the diagnosis of local tumour inoperability (defining T4), we do not expect that this will have a major impact on clinical management.[24]
In SPN, respiratory and cardiac motion can cause reduced ^{18}FDG PET signal (especially when located peripherally). A reduction of respiratory motion artefacts in PET imaging might be accomplished by respiratory gated ^{18}FDG PET-CT.[25] When using respiratory gating the partial volume effect will be less so that identification and perhaps quantification (radiotracer uptake e.g. SUV) of SPN might improve.[26] Gated ^{18}FDG PET-CT in radiotherapy planning could improve the definition of the tumour used for radiation treatment planning and, therefore increase the sparing of normal tissues resulting in a reduction of side effects.

Several studies showed significant improvement of accuracy in identifying nodal metastasis using combined ^{18}FDG PET and CT or hybrid ^{18}FDG PET-CT.[27-31] Problems in repositioning and movement artefacts are minimised. Finally, better anatomical orientation will definitely improve interobserver agreement of ^{18}FDG PET readings, which can be considerable in part as a function of experience [Smulders, submitted]. With the advent of alternative methods of invasive nodal staging (EUS, EBUS), often geographically complementary to surgical staging, and their likely impact on management of locally advanced IIIA patients prior and after induction therapy, there will be an increasing demand on imaging to provide better anatomical data.
Detection of unsuspected extrathoracic metastases has already been demonstrated by several studies in whole-body PET imaging.[2;6;29;32] The value of combined ^{18}FDG PET-CT in distant metastases could be locating, especially in unclear single focal abnormalities on ^{18}FDG PET. However the real advantage of ^{18}FDG PET-CT versus ^{18}FDG PET alone in distant staging needs to be verified.

It needs to be stressed that several technical issues with PET-CT have not been fully elucidated, that studies providing first-class evidence of superiority are scarce, and in this respect 'l'histoire se repète': just as with CT, MRI and PET technology rapidly diffuses into clinical practice apparently only slowed down by financial constraints. For those left with a recently purchased PET scanner, the question remains quoting Markus Schwaiger at the EANM congress in 2004: 'why with the advent of new technology, the old technique suddenly appears so very poor'.

Reference List

1. Van Tinteren H, Hoekstra O, Smit E et al. on behalf of the IKA-PLUS Study Group. Towards less Futile Surgery in Non Small Cell Lung Cancer? A Randomized Clinical Trial to Evaluate the Cost-effectiveness of Positron Emission Tomography. Controlled Clinical Trials 2001;22:89-98
2. Van Tinteren H, Hoekstra OS, Smit EF et al. Effectiveness of positron emission tomography in the preoperative assessment of patients with suspected non-small-cell lung cancer: the PLUS multicentre randomised trial. Lancet 2002;359:1388-1393
3. Zandwijk, N, Ruiz van Haperen, VWT, Smit, EF et al. Richtlijn voor toepassing van positron emissie tomografie met 18-Fluorodeoxyglucose (FDG PET) bij de diagnostiek van patiënten met niet-kleincellig bronchus carcinoom. IKA. 2001.
4. Visser O. Number of resections in non-small-cell lung cancer 1993-2002. 2005. Personal Communication
5. van Meerbeeck JP, Koning CC, Tjan-Heijnen VC et al. Guideline on 'non-small cell lung carcinoma; staging and treatment'. Richtlijn 'Niet-kleincellig longcarcinoom; stadiering en behandeling'. 2005;149:72-7
6. Pieterman RM, van Putten JW, Meuzelaar JJ et al. Preoperative staging of non-small-cell lung cancer with positron-emission tomography. N Engl J Med 2000;343:254-261
7. Vansteenkiste JF, Stroobants SG, De Leyn PR et al. Mediastinal lymph node staging with FDG-PET scan in patients with potentially operable non-small cell lung cancer: a prospective analysis of 50 cases. Leuven Lung Cancer Group. Chest 1997;112:1480-1486
8. Henschke CI, McCauley DI, Yankelevitz DF et al. Early Lung Cancer Action Project: overall design and findings from baseline screening. Lancet 1999;354:99-105
9. Kaneko M, Eguchi K, Ohmatsu H et al. Peripheral lung cancer: screening and detection with low-dose spiral CT versus radiography. Radiology 201(3), 798-802. 1996.
10. Sone S, Takashima S, Li F et al. Mass screening for lung cancer with mobile spiral computed tomography scanner. Lancet 351(9111), 1242-5. 2005.
11. Patz EF, Jr., Rossi S, Harpole DH et al. Correlation of tumor size and survival in patients with stage IA non- small cell lung cancer. Chest 2000;117:1568-1571
12. van Rens M.T, de la Riviere AB, Elbers HR et al. Prognostic assessment of 2,361 patients who underwent pulmonary resection for non-small cell lung cancer, stage I, II, and IIIA. Chest 2000;117:374-379
13. Lacasse Y, Wong E, Guyatt GH et al. Transthoracic needle aspiration biopsy for the diagnosis of localised pulmonary lesions: a meta-analysis. Thorax 1999;54:884-893
14. Herder GJ, Verboom P, Smit EF et al. Practice, efficacy and cost of staging suspected non-small cell lung cancer: a retrospective study in two Dutch hospitals. Thorax 2002;57:11-14
15. Fischer BM, Mortensen J, Hojgaard L. Positron emission tomography in the diagnosis and staging of lung cancer: a systematic, quantitative review. Lancet Oncol 2001;2:659-666
16. Gould MK, Maclean CC, Kuschner WG et al. Accuracy of positron emission tomography for diagnosis of pulmonary nodules and mass lesions: a meta-analysis. JAMA 2001;285:914-924

17. Higashi K, Ueda Y, Seki H et al. Fluorine- 18-FDG PET imaging is negative in bronchioloalveolar lung carcinoma. J Nucl Med 1999;39:1016-1020
18. Erasmus JJ, McAdams HP, Patz EFJ et al. Evaluation of primary pulmonary carcinoid tumors using FDG PET. AJR Am J Roentgenol 1998;170:1369-1373
19. Swensen SJ, Jett JR, Sloan JA et al. Screening for lung cancer with low-dose spiral computed tomography. Am J Respir Crit Care Med 2002;165:508-513
20. Sobue T, Moriyama N, Kaneko M et al. Screening for lung cancer with low-dose helical computed tomography: anti-lung cancer association project. J Clin Oncol 2002;20:911-920
21. Bradley J, Thorstad WL, Miller TR et al. Impact of FDG-PET on radiation therapy volume delineation in non-small-cell lung cancer. Int J Radiat Oncol Biol Phys 2004;59:78-86
22. Czermin J, Schelbert H. PET/CT imaging: facts, opinions, hopes and questions. J Nucl Med 2004;45:1S-3S
23. Ciernik IF, Dizendorf E, Baumert BG et al. Radiation treatment planning with an integrated positron emission and computer tomography (PET/CT): a feasibility study. Int J Radiat Oncol Biol Phys 2003;57:853-63
24. Comans EF. Correspondence. N Engl J Med 249, 1188-1190. 2003.
25. Nehmeh SA, Erdi YE, Rosenzweig KE et al. Reduction of respiratory motion artifacts in PET imaging of lung cancer by respiratory correlated dynamic PET: methodology and comparison with respiratory gated PET. J Nucl Med 2003;44:1644-8
26. Erdi YE, Nehmeh SA, Pan T et al. The CT motion quantitation of lung lesions and its impact on PET-measured SUVs. J Nucl Med 2004;45:1287-92
27. Aquino SL, Asmuth JC, Moore RH et al. Improved image interpretation with registered thoracic CT and positron emission tomography data sets. AJR Am J Roentgenol 2002;178:939-44
28. Cerfolio RJ, Ojha B, Bryant AS et al. The accuracy of integrated PET-CT compared with dedicated PET alone for the staging of patients with non small cell lung cancer. Ann Thor Surg 2004;78:1017-23
29. Lardinois D, Weder W, Hany TF et al. Staging of non-small-cell lung cancer with integrated positron-emission tomography and computed tomography. N Engl J Med 2003;348:2500-2507
30. Vansteenkiste JF, Stroobants SG, Dupont PJ et al. FDG-PET scan in potentially operable non-small cell lung cancer: do anatometabolic PET-CT fusion images improve the localisation of regional lymph node metastases? The Leuven Lung Cancer Group. Eur J Nucl Med 1998;25:1495-1501
31. Magnani P, Carretta A, Rizzo G et al. FDG/PET and spiral CT image fusion for mediastinal lymph node assessment of non-small cell lung cancer patients. J Cardiovasc Surg 1999;40:741-8
32. Weder W, Schmid RA, Bruchhaus H et al. Detection of extrathoracic metastases by positron emission tomography in lung cancer. Ann Thor Surg 1998;66:886-892

Het gebruik van ^{18}FDG PET bij het niet-kleincellig longcarcinoom

Hoofdstuk 1 bevat een inleiding van dit proefschrift. In Nederland is longkanker nog steeds het meest voorkomende type kanker bij mannen. De incidentie van longkanker voor mannen is 79.8 en voor vrouwen 27 per 100.000 inwoners per jaar (op basis van getallen uit 2000). De gemiddelde vijf-jaarsoverleving voor longkanker in Nederland is 10%, en hangt af van de histologische classificatie en het stadium. Bij ongeveer 84% gaat het om een niet-kleincellig longcarcinoom (NSCLC). De behandeling van NSCLC wordt bepaald door het stadium van de ziekte op het moment van presentatie. Het vaststellen van het juiste stadium is afhankelijk van de kwaliteit van verschillende diagnostische technieken. Positron emission tomography (PET), waarbij gebruik gemaakt wordt van ^{18}F-fluorodeoxyglucose (^{18}FDG), is een relatief nieuwe beeldvormende techniek. De accuratesse van ^{18}FDG PET bij NSCLC en de invloed op het patiëntenbeleid en het uiteindelijke resultaat is onderwerp voor veel studies.

Het werk dat hier gepresenteerd wordt, is een integraal deel van het onderzoek naar de rol van ^{18}FDG PET bij NSCLC zoals uitgevoerd in het VU medisch centrum. Dit onderzoek werd verricht in nauwe samenwerking met andere nationale en internationale universiteitscentra en perifere ziekenhuizen. De wetenschappelijk benadering in dit proefschrift volgt het raamwerk ter bestudering van de effectiviteit van diagnostische testen zoals beschreven door van Tinteren et al.[van Tinteren, submitted]. Dit raamwerk begint met een beoordeling van de huidige praktijk. Daarbij worden eventuele inefficiënties van de huidige praktijk aangetoond. De volgende stap is analyse van de beschikbare literatuur en beoordeling van de data ten aanzien van accuratesse van de nieuwe test. Indien deze onvoldoende zijn, dan moeten deze testen alsnog verricht worden. Indien er voornamelijk accuratesse data van de nieuwe test beschikbaar zijn, dan is de volgende stap het beoordelen van potentiële invloed van de toegevoegde nieuwe test aan de diagnostische strategie (klinische waarde). Tegelijkertijd worden er modelmatige studies verricht ter beoordeling van kosten-effectieve toepassingen van de nieuwe test, welke dan getoetst kunnen worden in een gerandomiseerde studie. Uiteindelijk zou dit moeten leiden tot de ontwikkeling en implementatie van "evidence based"-richtlijnen. Na de richtlijn beoordeeld te hebben zou een nieuwe evaluatie opgestart kunnen worden waarbij de nieuwste trends (nieuwe technieken of strategieën) ook geëvalueerd worden.
Het bovenstaande schema is toegepast bij de stadiëring van patiënten met een verdenking op NSCLC, daarbij zijn studies verricht die in de diverse hoofdstukken worden besproken:
Hoofdstuk 2 beschrijft de eerste stap in het raamwerk waarin de klinische praktijk beschreven wordt, opbrengst en kosten van preoperatieve stadiëring bij patiënten met (verdacht) NSCLC in een academisch en algemeen ziekenhuis in een tijdsbestek van 2 jaar (1993/1994). Chirurgische behandeling bleek onvoldoende bij bijna 50% van de geopereerde patiënten met NSCLC, ten gevolge van een niet resectabele tumor of een benigne afwijking tijdens operatie of een metastase dan wel recidief binnen 1 jaar na operatie. Tijdens operatie bleken 33 niet resectabel (23% van de patiënten die een operatie ondergingen), 19 (13%) hadden een benigne afwijking. Bij 18 patiënten werd de chirurgische behandeling beschouwd als

futiel vanwege een metastase (13%), en bij 1 patiënt (0.7%) in verband met lokaal recidief binnen 12 maanden na een in opzet curatieve chirurgische behandeling. Er waren gemiddeld vijf (standaard deviatie ± 1.5) testen nodig om tot een uiteindelijke stadiëring te komen, in een tijdsbestek van 20 dagen (mediaan), bij 13% van de patiënten was meer dan 6 weken nodig. Bij de meeste patiënten blijkt het aantal testen, de duur en de kosten nodig ter stadiëring aanzienlijk te zijn. Het resultaat van de stadiëring, met als doel het voorkomen van onnodige operaties, is matig.

Zoals eerder gezegd, betrof 1 van de 7 ten onrechte uitgevoerde operaties patiënten welke uiteindelijk geen maligniteit bleken te hebben. In de literatuur is veel geschreven over de accuratesse van ^{18}FDG PET bij radiologisch onduidelijke longhaarden, er is echter zeer weinig geschreven over de accuratesse van ^{18}FDG PET bij kleine longhaarden. Hoofdstuk 3 beschrijft de karakteristieken van ^{18}FDG PET als een functie van pre-test risico-inschatting bij radiologisch onduidelijke solitaire longhaarden ≤10 mm op spiraal CT-scan.
In deze retrospectieve studie hebben wij 35 patiënten geïncludeerd met 36 solitaire long haarden ≤10 mm (uitersten: 3-10 mm) in diameter bij klinische presentatie. ^{18}FDG PET identificeerde 30 van 36 kleine laesies correct. Een afwijking bleek fout negatief op^{18}FDG PET en bij vijf laesies bleek ^{18}FDG PET fout positief. Specificiteit was 77% (17/22;95% betrouwbaarheidsinterval:0.55-0.92), sensitiviteit was 93% (13/14; 95% betrouwbaarheidsinterval:0.66-1.0), de positief voorspellende waarde was 72% (13/18;95% betrouwbaarheidsinterval:0.46-0.90) en de negatief voorspellende waarde 94% (17/18; 95% betrouwbaarheidsinterval:0.73-1.0). Deze data suggereren dat ^{18}FDG PET een waardevolle diagnostische test is ter differentiatie (benigne versus maligne) van onduidelijke solitaire longhaarden ≤10 mm in diameter op spiraal CT scan bij presentatie.

In de grote hoeveelheid ^{18}FDG PET -literatuur werd overvloedig de accuratesse van ^{18}FDG PET bij longhaarden beschreven en verrassend weinig over de toegevoegde waarde van ^{18}FDG PET versus een klinische kansschatting ondanks dat dit van belang is voor een verbetering van diagnostisch begrip en de invloed op therapiekeuze.
De meeste clinici zullen intuïtief een kansschatting maken, hoewel de variabiliteit tussen clinici onderling vooralsnog onbekend is, lijkt een objectiever scoringssysteem om de kans op maligniteit in te schatten te prefereren. Swensen et al. hebben een voorstel gedaan voor een intern gevalideerde wiskundige formule waarin een kansschatting op maligniteit weergegeven wordt als een functie van 3 klinische (leeftijd, roken, voorgeschiedenis met kanker) en 3 radiologische (diameter, uitlopers, locatie) variabelen. In hoofdstuk 4 beschrijven we een externe validatie van het voorspellingsmodel waarin we gebruik maken van CT en ^{18}FDG PET-data van 106 patiënten welke verwezen werden in verband met onduidelijkheid over de aard van de longhaard. Tegelijkertijd hebben wij de toegevoegde waarde van ^{18}FDG PET gekwantificeerd als een functie van zijn operationele karakteristieken in relatie tot dit

predictiemodel. ROC-analyse suggereert dat de beste resultaten te verwachten zijn van klinische beoordeling en ^{18}FDG PET gecombineerd.
Ter beoordeling van de potentiële klinische waarde van ^{18}FDG PET bij patiënten met verdachte longhaarden met een diagnostisch probleem, hebben wij een voor/na-studie ontworpen (hoofdstuk 5). Dit soort studies geeft informatie over het diagnostisch begrip en de invloed op therapeutisch beslissingen. Het prospectieve onderzoek bestond uit het gebruik van vragenlijsten voor, meteen na en enkele maanden na de resultaten van ^{18}FDG PET. Patiënten werden met name verwezen naar het ^{18}FDG PET centrum voor onduidelijke radiologische longhaarden, mediastinale stadiëring of onduidelijke bevindingen bij conventionele stadiëring ten aanzien van beoordeling van metastasen op afstand. Een toegenomen diagnostisch begrip werd bij 84% gerapporteerd en bij 50% van de patiënten vond een verandering van het therapeutische beleid plaats. Het afzeggen van een operatie werd het meest frequent gerapporteerd (35%). In de groep patiënten welke verwezen werden vanwege onduidelijke radiologische bevindingen, vond een verbetering van het diagnostische begrip bij 46% plaats en een zinvolle verandering van therapie bij 71% van de patiënten.

Hoofdstuk 6 beschrijft een beslissingsmodel ter beoordeling van wanneer en hoe ^{18}FDG PET kosten-effectief zou kunnen zijn bij routineus gebruik van ^{18}FDG PET in de preoperatieve stadiëring van NSCLC. Praktijkvariatie werd gevonden tussen 2 ziekenhuizen met betrekking tot locatie waar de diagnostiek plaats vond (klinisch versus poliklinisch) en het gebruik van mediastinoscopie. Dit kwam terug in de kosten en in het aantal operaties. Hospitalisatie was de grootste kostenpost bij deze patiëntenpopulatie. Introductie van ^{18}FDG PET voor alle patiënten zal leiden tot een substantiële toename van de stadiëringskosten, welke voor een deel "terugverdiend" worden door een daling van het aantal ten onrechte uitgevoerde operaties. ^{18}FDG PET uitgevoerd voor invasieve stadiëring zal leiden tot een beperkte toename van kosten. Deze beperking komt met name door substitutie van diagnostische testen. Er zal altijd een toename in kosten plaatsvinden wanneer ^{18}FDG PET alleen gebruikt zal worden voor patiënten met NSCLC welke een operatie kunnen ondergaan. Met name door vermindering van het aantal ten onrechte uitgevoerde operaties kunnen de kosten beperkt worden. Vanuit het perspectief van kosten is het gebruikt van ^{18}FDG PET bij NSCLC het meest optimaal na beeldvormende diagnostiek en voor de invasieve stadiëring.

De meest betrouwbare en overtuigende informatie ten aanzien van de rol van ^{18}FDG PET in het stadiëringstraject van NSCLC is gerandomiseerd onderzoek.[1] Wij hebben een gerandomiseerde trial [2] uitgevoerd ter bestudering van de toegevoegde waarde van ^{18}FDG PET net voor invasief mediastinaal onderzoek of thoracotomie en een tweede trial (beschreven in hoofdstuk 7) ter beoordeling van de rol van ^{18}FDG PET vroeg in het diagnostische traject gepositioneerd. Daarbij werden ook de substitutiemogelijkheden van ^{18}FDG PET beoordeeld.
De onderzoeksvraag was of de toepassing van ^{18}FDG PET meteen na eerste presentatie, het stadiëringstraject zou kunnen vereenvoudigen zonder aan accuratesse te verliezen. Controle-

arm was de standaard (conventionele) stadiëringsstrategie (volgens internationale richtlijnen) zonder ^{18}FDG PET. Patiënten werden gerandomiseerd direct na eerste verdenking op NSCLC (patiënten met bij eerste presentatie duidelijk gemetastaseerde ziekte werden uitgesloten), vervolgens volgde histologische of cytologisch verificatie van afwijkingen of beeldvormende diagnostiek en follow-up.
Tussen 1999 en 2001 werden 465 patiënten (233 conventionele arm, 232 PET arm) geïncludeerd door 22 ziekenhuizen. Het gemiddelde (standaard deviatie) aantal verrichtingen wat nodig was voor stadiëring in de 2 armen (conventioneel en PET) was 7.9 (2.0) vs. 7.9 (1.9); p=0.90. Overeenkomst tussen klinische stadiëring en de stadiëring na operatie of FU was goed in beide armen (Kappa 0.85 vs. 0.78; p=0.07). De kosten verschilden niet significant. We kunnen concluderen dat het gebruik van ^{18}FDG PET vroeg in het stadiërings–traject bij patiënten met (verdenking op) NSCLC gelijke kwaliteit heeft vergeleken met de conventionele arm, en geen vereenvoudiging van stadiëring geeft.

Richtlijnontwikkeling

In 2001 werd een voorlopige richtlijn voor het gebruik van ^{18}FDG PET bij NSCLC samengesteld door een regionale multidisciplinaire commissie.[3] Samen met de groeiende beschikbaarheid van ^{18}FDG PET in onze regio, resulteerde dit in een betere toegang tot ^{18}FDG PET voor deze indicatie. Data vanuit het Integraal Kanker Centrum Amsterdam lieten een blijvende daling zien van het aantal longresecties in de orde van grootte zoals voorspeld in de eerste gerandomiseerde trial.[4]

In 2004 werden "evidence based"richtlijnen voor het diagnostische traject en behandeling van NSCLC ontwikkeld door een nationaal multidisciplinaire commissie.[5] ^{18}FDG PET wordt geadviseerd bij patiënten met stadium I/II/III NSCLC die in aanmerking komen voor curatieve therapie, na conventionele work-up en voor invasieve mediastinale stadiëring. Bij verdenking op locoregionale of haematogene metastasen wordt verificatie geadviseerd. Mediastinoscopie of endobronchiale echo (EBUS) of transoesophageale echo (EUS) wordt geadviseerd indien de primaire tumor tegen het mediastinum aan ligt op ^{18}FDG PET. Door de resolutie van ^{18}FDG PET kan geen onderscheid gemaakt worden tussen aanpalende klieren of primaire tumor bij centrale tegen het mediastinum aanliggende tumoren.[6,7] Bij tumoren die ^{18}FDG opnemen, zonder dat er aanwijzingen zijn voor mediastinale dan wel afstandsmetastasen, wordt chirurgie geadviseerd. Op moment van schrijven wordt de nieuwe richtlijn geïmplementeerd. Na implementatie zal ook deze richtlijn getoetst moeten worden op tekortkomingen.

Klinisch en radiologisch onduidelijke coin-laesies komen niet aan bod in bovengenoemde richtlijn. Gebruik makend van de beschikbare literatuur, zouden wij de volgende benadering willen voorstellen.

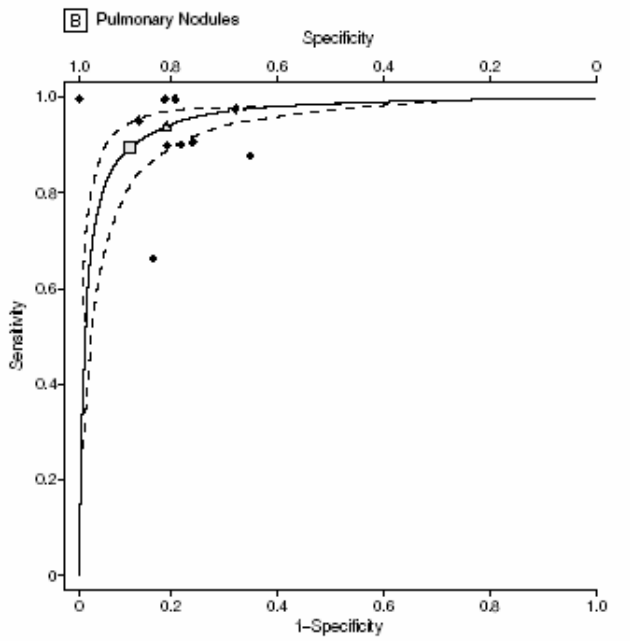
Het gebruik van ^{18}FDG PET bij klinisch en radiologisch onduidelijke coin-laesies ≥ 1 cm

Data van longkanker screening laten zien dat spiraal CT sensitiever is dan röntgenfoto's van de longen bij het opsporen van kleine afwijkingen.[8-10] Ten gevolge van deze screenings-studies en de verder gaande ontwikkeling van CT zullen er meer en kleinere longafwijkingen ontdekt worden. Het is van belang om zo snel mogelijk te beoordelen of zo'n coin-laesie maligne is, gezien de goede vijf-jaarsoverleving jaars overleving bij patiënten met een stadium IA NSCLC (80% [11;12]). De volgende stap in het diagnostische traject moet duidelijk zijn. Verschillende strategieën kunnen gevolgd worden zoals transthoracale naaldbiopsie, operatie, afwachtend beleid of ^{18}FDG PET. De keuze zal afhangen van de sensitiviteit en specificiteit van een test, morbiditeit en de kosten. De kans op een niet-diagnostisch resultaat bij een transthoracaal naaldbiopt is aanzienlijk bij afwijkingen kleiner dan 2 cm, het risico op een pneumothorax is ook aanwezig (3.1- 41.7% [13]).

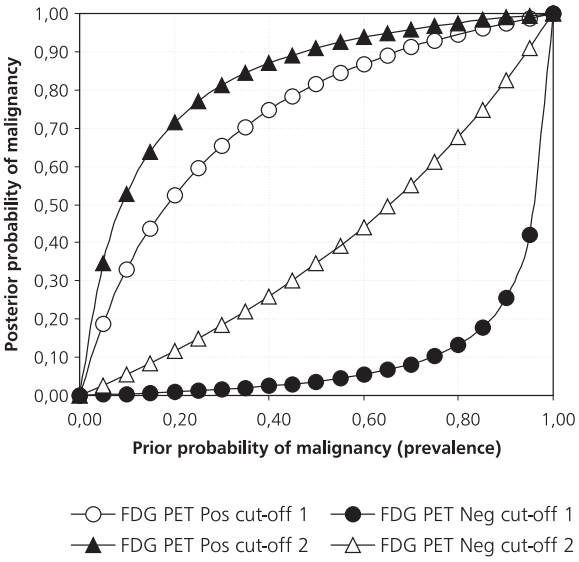
Zoals eerder genoemd in dit proefschrift had 15% van alle geopereerde patiënten een goedaardige afwijking.[14]

Afwachtend beleid draagt het risico met zich mee, dat het uiteindelijk toch om een maligne aandoening gaat, waarbij afwachten mogelijk ongunstig voor de behandeling zou kunnen zijn. Daarom is er behoefte aan accurate niet-invasieve testen ter voorkoming van chirurgische interventies voor benigne afwijkingen.

Selectie van chirurgische kandidaten zou verbeterd worden indien ^{18}FDG PET toegevoegd wordt aan het diagnostische traject. Samengevoegde sensitiviteit en specificiteit van ^{18}FDG PET bij solitaire longhaarden ≥ 1 cm was 97% en 78%.[15] Dit impliceert een kans op maligniteit van 1% wanneer de pre-test kansschatting 20% is en het ^{18}FDG PET resultaat negatief is versus een post-test kans op maligniteit van 86% wanneer de pre-test kans 80% is. Het gebruik van ^{18}FDG PET bij deze patiënten is gerelateerd aan de kans op verandering van het therapeutische plan. De keuze van behandeling is afhankelijk van de kans op maligniteit. Dit hangt af van klinische en radiologische bevindingen. Fischer liet eerder zien dat ^{18}FDG PET gebruikt moet worden in een populatie met een pre-test kans op kanker tussen de 10 en 50%.[15;16] Een negatieve ^{18}FDG PET-scan met een prevalentie van 50% of minder sluit een maligniteit uit. Wanneer ^{18}FDG PET gebruikt wordt in een populatie met een hogere prevalentie (meer dan 50%), dan zou een negatieve ^{18}FDG PET-scan een hoger post-test kans op maligniteit geven en moet aanvullend onderzoek verricht worden. Bij ^{18}FDG PET negatieve coin-laesies kan differentiaal diagnostisch gedacht worden aan bronchioalveolair cel carcinoma, carcinoid, metastase van Grawitz, neuro-endocriene tumoren etc.[17,18] Indien een afwachtend beleid aangehouden wordt, dan zouden radiologische testen herhaald



Figuur 1. De zwarte ruitjes laten de individuele studie resultaten zien (sensitiviteit en 1-specificiteit). Het grijze vierkantje geeft een indicatie van de maximum "joint" sensitiviteit en specificiteit (een globale maat voor accuratesse). Het grijze driehoekje is het Q-punt op de ROC-curve van ^{18}FDG PET in de huidige praktijk ten aanzien van het detecteren van maligne longhaarden. Figuur van Gould et al. Jama 2001.[16]



Figuur 2. Voorspelde post-test kansschatting van ^{18}FDG PET bij verschillende uitvoerings-karakteristieken (1 afkapwaarde volgens de standaard manier sensitief PET beoordelen: sensitiviteit 0.94 en specificiteit 0.77; 2 meer conservatieve interpretatie van afkapwaarde: sensitiviteit 0.50 en specificiteit 0.95).

moeten worden gedurende 18 maanden. Indien de afwijking groeit, moet gekozen worden voor een operatie (verificatie).
^{18}FDG PET bij coin-laesies die kleiner zijn dan 1 cm moet nog beter bestudeerd worden. Accuratesse zou eenvoudig beoordeeld kunnen worden bij een geselecteerde subset van patienten met onduidelijke radiologische bevindingen op CT. Voor de uiteindelijke uitvoeringskarakteristieken zijn ROC-analyses nodig. Het Q-punt (het punt op de curve welke het dichtste bij de linker bovenhoek ligt op de grafiek) bij Gould en anderen correspondeerde met een sensitiviteit van 94% en een specificiteit van 83% (Figuur 1).

Het is waarschijnlijk dat bij lage pre-test kansschatting een ^{18}FDG positieve PET-scan volgens de standaard criteria niet richting gevend is (Figuur 2). Deze voorspellingen zijn hypothetisch, ROC-curven zouden anders kunnen zijn voor kleine afwijkingen door 'partial volume' effect. Sinds kort is er ook de mogelijkheid tot 'gating' bij ^{18}FDG PET en ^{18}FDG PET-CT scanners. Studies ter beoordeling van accuratesse maar ook kwantificatie van ^{18}FDG opname in kleine laesies met de huidige PET-scanners en de nieuwere technieken (gated ^{18}FDG PET en ^{18}FDG PET-CT) zijn zeer wenselijk. Tegelijkertijd zou de maat van onzekerheid wat de patiënt maar ook de dokter zou willen/kunnen accepteren onderzocht moeten worden.

Toekomst

PET-CT

Een nieuwe ontwikkeling in de beeldvormende diagnostiek is het combineren van 2 of meer technieken zoals ^{18}FDG PET en CT. Tot nu toe werden deze 2 technieken apart uitgevoerd en beoordeeld. Sinds kort zijn er ook geïntegreerde ^{18}FDG PET-CT-systemen. Deze systemen produceren ^{18}FDG PET- en CT-beelden welke nagenoeg simultaan verkregen zijn, wat de interpretatie van de beelden vereenvoudigd en mogelijk ook verbeterd.
Het belangrijkste voordeel van ^{18}FDG PET-CT versus beide technieken separaat bij NSCLC is de mogelijkheid om onderscheid te maken tussen tumor en atelectase.[19] Hoewel er geluiden opgaan dat ^{18}FDG PET-CT bij kan dragen ten aanzien van het beter beoordelen van tumorgrootte (met name T4) verwachten wij niet dat dit een belangrijke invloed heeft op het klinische beleid.[20]
Bij solitaire longhaarden kan het ademen en de hart-activiteit een verminderd ^{18}FDG-signaal geven. Een vermindering van bewegings-artefacten door ademen zou bereikt kunnen worden door 'gated' ^{18}FDG PET-CT.[21] Het 'partial volume' effect zou minder kunnen zijn zodat identificatie en mogelijk ook kwantificatie van SPN zou kunnen verbeteren.[22] Het beter beoordelen van de tumorgrootte (T stadium) en minder bewegingartefact ten gevolge van 'gating' zou ook bij de radiotherapieplanning van voordeel kunnen zijn, waarbij meer normaal weefsel gespaard kan worden met daardoor minder bijwerkingen.

Verschillende studies laten een verbetering zijn van accuratesse ten aanzien van het identificeren van mediastinale metastasen bij het gebruik van ^{18}FDG PET-CT.[23-27] Uiteindelijk zal een betere anatomische oriëntatie leiden tot een verbetering van interobserver overeenkomst van het beoordelen van ^{18}FDG PET [Smulders et al, submitted].

Met de komst van nieuwe minder invasieve technieken ter beoordeling van mediastinale lymfklieren (EUS, EBUS) zal er ook meer vereist worden van beeldvormende technieken om betere anatomische data te verkrijgen. Mogelijk speelt ^{18}FDG PET-CT daar ook een rol bij.

De meerwaarde van ^{18}FDG PET-CT bij afstandsmetastasen zou kunnen liggen bij het beter lokaliseren, met name bij onduidelijke solitaire afwijkingen op afstand. Het echte voordeel van ^{18}FDG PET-CT ten opzichte van ^{18}FDG PET alleen bij het beoordelen van afstandsmetastasen moet nog bewezen worden.

Referentie Lijst

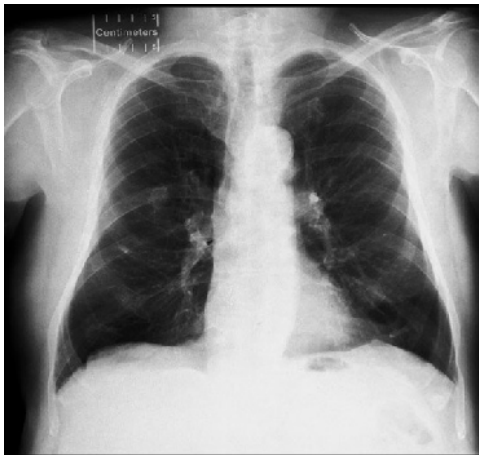
1. Van Tinteren H, Hoekstra O, Smit E et al. on behalf of the IKA-PLUS Study Group. Towards less Futile Surgery in Non Small Cell Lung Cancer? A Randomized Clinical Trial to Evaluate teh Cost-effectivensess of Positron Emission Tomography. Controlled Clinical Trials 2001;22:89-98
2. Van Tinteren H, Hoekstra OS, Smit EF et al. Effectiveness of positron emission tomography in the preoperative assessment of patients with suspected non-small-cell lung cancer: the PLUS multicentre randomised trial. Lancet 2002;359:1388-1393
3. Zandwijk, N, Ruiz van Haperen, VWT, Smit et al. Richtlijn voor toepassing van positron emissie tomografie met 18F-Fluorodeoxyglucose (FDG PET) bij de diagnostiek van patienten met niet-kleincellig bronchus carcinoom. IKA. 2001.
4. Visser O. Number of resections in non-small-cell lung cancer 1993-2002. 2005.
5. van Meerbeeck JP, Koning CC, Tjan-Heijnen VC et al. [Guideline on 'non-small cell lung carcinoma; staging and treatment']. Richtlijn 'Niet-kleincellig longcarcinoom; stadiëring en behandeling'. 2005;149:72-7
6. Pieterman RM, van Putten JW, Meuzelaar JJ et al. Preoperative staging of non-small-cell lung cancer with positron- emission tomography. N Engl J Med 2000;343:254-261
7. Vansteenkiste JF, Stroobants SG, De Leyn PR et al. Mediastinal lymph node staging with FDG-PET scan in patients with potentially operable non-small cell lung cancer: a prospective analysis of 50 cases. Leuven Lung Cancer Group. Chest 1997;112:1480-1486
8. Henschke C I, McCauley DI, Yankelevitz DF et al. Early Lung Cancer Action Project: overall design and findings from baseline screening. Lancet 1999;354:99-105
9. Kaneko, M, Eguchi, K, Ohmatsu, H et al. Peripheral lung cancer: screening and detection with low-dose spiral CT versus radiography. Radiology 201(3), 798-802. 1996. Ref Type: Journal (Full)
10. Sone S, Takashima S, Li F et al. Mass screening for lung cancer with mobile spiral computed tomography scanner. Lancet 351(9111), 1242-5. 2005. Ref Type: Journal (Full)
11. Patz EF, Jr, Rossi S, Harpole DH, Jr. et al. Correlation of tumor size and survival in patients with stage IA non- small cell lung cancer. Chest 2000;117:1568-1571
12. van Rens MT, de la Riviere AB, Elbers HR et al. Prognostic assessment of 2,361 patients who underwent pulmonary resection for non-small cell lung cancer, stage I, II, and IIIA. Chest 2000;117:374-379
13. Lacasse Y., Wong E, Guyatt GH et al. Transthoracic needle aspiration biopsy for the diagnosis of localised pulmonary lesions: a meta-analysis. Thorax 1999;54:884-893
14. Herder GJ, Verboom P, Smit EF et al. Practice, efficacy and cost of staging suspected non-small cell lung cancer: a retrospective study in two Dutch hospitals. Thorax 2002;57:11-14

15. Fischer BM, Mortensen J, Hojgaard L. Positron emission tomography in the diagnosis and staging of lung cancer: a systematic, quantitative review. Lancet Oncol. 2001;2:659-666
16. Gould MK, Maclean CC, Kuschner WG et al. Accuracy of positron emission tomography for diagnosis of pulmonary nodules and mass lesions: a meta-analysis. JAMA 2001;285:914-924
17. Higashi K, Ueda Y, Seki H et al. Fluorine- 18-FDG PET imaging is negative in bronchioloalveolar lung carcinoma. Journal of Nuclear Medicine 1999;39:1016-1020
18. Erasmus JJ, McAdams HP, Patz EFJ et al. Evaluation of primary pulmonary carcinoid tumors using FDG PET. AJR. 1994;1369-1373
19. Bradley J, Thorstad WL, Miller TR et al. Impact of FDG-PET on radiation therapy volume delineation in non-small-cell lung cancer. Int J Radiat Oncol Biol Phys 2004;59:78-86
20. Comans EF. Correspondence. N.Engl.J.Med. 249, 1188-1190. 2003.
21. Nehmeh SA, Erdi YE, Rosenzweig KE et al. Reduction of respiratory motion artifacts in PET imaging of lung cancer by respiratory correlated dynamic PET: methodology and comparison with respiratory gated PET. J Nucl Med 2003;44:1644-8
22. Erdi YE, Nehmeh SA, Pan T et al. The CT motion quantitation of lung lesions and its impact on PET-measured SUVs. J Nucl Med 2004;45:1287-92
23. Aquino SL, Asmuth JC, Moore RH et al. Improved image interpretation with registered thoracic CT and positron emission tomography data sets. AJR Am J Roentgenol 2002;178:939-44
24. Cerfolio RJ, Ojha B, Bryant AS et al. The accuracy of integrated PET-CT compared with dedicated pet alone for the staging of patients with nonsmall cell lung cancer. Ann Thorac Surg 2004;78:1017-23
25. Lardinois D, Weder W, Hany TF et al. Staging of non-small-cell lung cancer with integrated positron-emission tomography and computed tomography. N Engl J Med 2003;348:2500-2507
26. Vansteenkiste JF, Stroobants SG, Dupont PJ et al. FDG-PET scan in potentially operable non-small cell lung cancer: do anatometabolic PET-CT fusion images improve the localisation of regional lymph node metastases? The Leuven Lung Cancer Group. European Journal of Nuclear Medicine 1998;25:1495-1501
27. Magnani P, Carretta A, Rizzo G et al. FDG/PET and spiral CT image fusion for medistinal lymph node assessment of non-small cell lung cancer patients. J Cardiovasc Surg 1999;40:741-8

Met andere woorden...

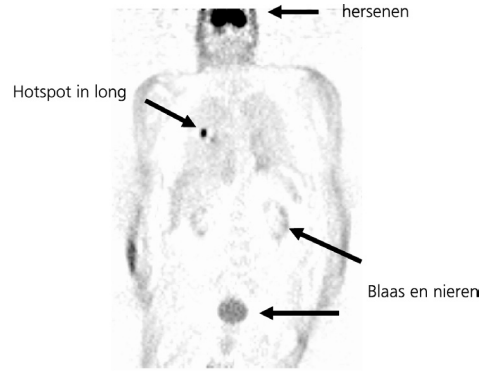
Longkanker is nog steeds een van de meest voorkomende vormen van kanker. Per jaar wordt in Nederland bij ruim 8000 nieuwe patiënten longkanker gediagnosticeerd.[1] Er zijn verschillende typen longkanker. Verreweg het grootste aandeel (84%) van longkanker wordt histologisch geclassificeerd als niet-kleincellig longkanker. De 5-jaars overleving voor niet-kleincellig longkanker onafhankelijk van het stadium van voortschrijding van de ziekte is 10%. Het hoge aantal sterfgevallen wordt veroorzaakt doordat longkanker heel vaak pas in een gevorderd stadium wordt ontdekt waardoor een genezende behandeling niet meer mogelijk is. De behandeling en prognose van niet-kleincellig longkanker zijn afhankelijk van het stadium, dat wil zeggen de mate van uitzaaiingen. Hierbij gaat het om de grootte en uitbreiding van de tumor (lokaal), uitzaaiingen naar lymfklieren en/of uitzaaiingen naar andere organen in het lichaam (bijvoorbeeld bot of lever). Om het stadium te bepalen zijn bepaalde onderzoeken nodig: dit kunnen beeldvormende technieken zijn zoals een röntgenfoto, computertomografie (Figuur 1) of meer invasief zoals het verkrijgen van weefsel door middel van bijvoorbeeld een operatie. Aan de hand van deze onderzoeken wordt bepaald of een patiënt in aanmerking komt voor een operatie, radiotherapie, chemotherapie of een gecombineerde behandeling (bijvoorbeeld chemotherapie en een operatie).
In *hoofdstuk 2* wordt dit proces van stadiering beschreven in twee ziekenhuizen tussen 1993 en 1994. Deze studie laat zien dat het traject langdurig en ingewikkeld kan zijn met een niet-optimaal resultaat. Er waren in dit onderzoek gemiddeld vijf testen (in ongeveer 20 dagen) nodig voor het bepalen van het stadium. Uiteindelijk blijkt dat de helft van de patiënten ten onrechte geopereerd is. Tijdens de operatie bleek bij 23% van de patiënten de tumor niet volledig verwijderd te kunnen worden, bij 13% van de patiënten er een goedaardige afwijking te zijn en werd bij 13% van de patiënten binnen 12 maanden na de operatie een uitzaaiing gevonden. Bij 50% van de geopereerde patiënten bleek het uiteindelijke stadium (bij operatie en follow-up) niet overeen te komen met het stadium bepaald voorafgaand aan de operatie.

Sinds een aantal jaren is er een nieuwe beeldvormende techniek, ^{18}F-fluorodeoxyglucose positron emissie tomografie (^{18}FDG PET), waarbij met name de stofwisselingsprocessen zichtbaar gemaakt worden, dit in tegenstelling tot bijvoorbeeld computertomografie waarbij informatie verkregen wordt over grootte en vorm van de afwijkingen.
Kankercellen hebben een verhoogde stofwisseling en nemen onder andere meer glucose op dan andere cellen. Door glucose te koppelen aan een radioactief fluoratoom (^{18}F) wordt een radioactieve verbinding verkregen: ^{18}FDG. Na injectie via een ader en verspreiding in het lichaam via de bloedbaan, zal dit radioactieve glucose zich vooral in een orgaan of weefsel met een verhoogde stofwisseling ophopen. Door middel van een PET-camera kan deze verhoogde stofwisseling (“hotspot”) in beeld gebracht worden. Daarbij wordt het hele lichaam gescand



Figuur 1. A. Röntgen foto van de longen. Geen duidelijke afwijking zichtbaar.

Figuur 1. B. Computer tomografie van de longen. Pijl: verdachte afwijking voor longkanker in longweefsel.



Figuur 1. C. ^{18}FDG PET scan, normale opname in hersenen, nieren en blaas. Hotspot in rechter long.

in tegenstelling tot andere beeldvormende technieken waarbij doorgaans maar een deel van het lichaam wordt onderzocht (Figuur 1). Eerder onderzoek laat zien dat een ^{18}FDG PET-scan beter zou kunnen zijn voor het in kaart brengen van het stadium bij niet-kleincellig longkanker. Omdat de ^{18}FDG PET-scan ook niet perfect is, ook ontstekingen kunnen een hotspot geven, moeten de hotspots bevestigd worden door microscopisch weefselonderzoek.

Door verbetering van technieken en breder inzetten van bepaalde technieken (bijvoorbeeld computer tomografie als routineonderzoek van bevolkingsgroepen met een verhoogd risico op ontwikkeling van longkanker) zullen er ook meer longafwijkingen gevonden worden waarvan op radiologisch criteria het onduidelijk is wat de aard van deze afwijkingen is. Deze radiologisch niet te classificeren longafwijkingen kunnen ook modelmatig bekeken worden. *Hoofdstuk 3* laat een model zien van een kansschatting naar kwaadaardigheid van onduidelijke longafwijkingen (≤ 3 cm) op computertomografie. Hierbij wordt gebruikt gemaakt van bepaalde kenmerken van longkanker op computer tomografie (locatie, vorm, diameter van afwijking) en klinische kenmerken (leeftijd, wel of niet roken, wel of geen kanker gehad). Dit reeds eerder beschreven model combineren we met ^{18}FDG PET. Dit onderzoek laat zien dat de combinatie van de klinische en radiologische kenmerken en het resultaat van ^{18}FDG PET een goede combinatie is voor de voorspelling (goed- of kwaadaardig) van onduidelijke longafwijkingen.

Naast een toename in het aantal radiologisch moeilijk te classificeren longhaarden is het ook mogelijk om steeds kleinere tumoren te ontdekken. Het beoordelen van kleine perifere afwijkingen in de longen is eveneens een moeilijk proces. Bij computer tomografie wordt zo goed mogelijk een beoordeling gegeven (kwaadaardig, goedaardig, onduidelijk) aan de hand van verschillende criteria. Met name de onduidelijke longafwijkingen geven vaak aanleiding tot veel onderzoek en soms ook operaties welke achteraf niet nodig blijken te zijn. *Hoofdstuk 4* beschrijft de waarde van ^{18}FDG PET bij deze kleine (≤10 mm) op computertomografie niet te classificeren afwijkingen. ^{18}FDG PET kon in 93% van de afwijkingen de afwijking ook daadwerkelijk zichtbaar maken en in 77% van de afwijkingen was deze afwijking correct beoordeeld als goed- of kwaadaardig. Ook bij deze kleine (≤10 mm) afwijkingen lijkt ^{18}FDG PET zeker van aanvullende waarde.

Wanneer nieuwe technieken gebruikt worden is het ook belangrijk dat ze beoordeeld worden in de praktijk. Wat zijn bijvoorbeeld de redenen voor ^{18}FDG PET-gebruik, hoe nuttig vond de aanvrager deze ^{18}FDG PET-scan en is er wat veranderd in het beleid? Een manier om dit te beoordelen is door middel van vragenlijsten voor, direct na en ruim na het maken van een ^{18}FDG PET-scan. In *hoofdstuk 5* beschrijven we deze aanpak en de resultaten hiervan bij de evaluatie van longkanker. Patiënten in dit onderzoek werden met name verwezen naar het klinisch PET-centrum vanwege onduidelijke bevindingen bij beeldvormend onderzoek. Artsen rapporteerden in 84% een beter begrip en er vond in 59%, naar aanleiding van het PET-resultaat, een gunstige beleidsverandering plaats. De belangrijkste therapeutische consequentie naar aanleiding van de uitslag van de ^{18}FDG PET-scan was het afzeggen van een operatie (in 35% van de patiënten).

^{18}FDG PET-scan is kostbaar en niet overal vrij beschikbaar. Gezien de beperkte capaciteit in Nederland moet bepaald worden voor welke groep patiënten deze onderzoekstechniek zinvol is. Aan de hand van eerder onderzoek werd een model ontwikkeld om de invloed van ^{18}FDG PET te bepalen op de stadiering bij niet-kleincellig longkanker en zicht te krijgen op de kosten. *Hoofdstuk 6* laat op een modelmatige manier zien dat ^{18}FDG PET bij de stadiering van niet-kleincellig longkanker kosteneffectief is na beeldvormend onderzoek en voor invasief onderzoek. Het nadeel van een kosten-effectiviteits studie is, is dat deze gebaseerd is op een aantal aannames. Een manier om dit te omzeilen is een vergelijkende studie waarbij de ene groep patiënten een PET scan krijgt en de andere groep geen PET scan.

Het laatste hoofdstuk (*hoofdstuk 7*) in dit proefschrift beschrijft de studie waarbij gekeken is of het stadiëringstraject bij patiënten met (verdenking op) niet-kleincellig longkanker eenvoudiger gemaakt kan worden (bijvoorbeeld weglaten van bepaalde onderzoeken) zonder minder nauwkeurig te worden. Hiervoor is een vergelijkende studie gedaan tussen de "traditionele" stadiering volgens de internationale richtlijnen (dit is zonder ^{18}FDG PET-scan) en de stadiering waarbij gebruik gemaakt werd van ^{18}FDG PET-scan direct nadat bekend werd

dat er een verdachte afwijking op de foto zichtbaar was. Het aantal onderzoeken dat nodig was voor het bepalen van het stadium en de kosten van onderzoek en behandeling, was in beide groepen gelijk. De duur van het onderzoekstraject was beduidend korter in de ^{18}FDG PET-groep. Beide groepen waren even nauwkeurig ten aanzien van het juist bepalen van het stadium. We kunnen stellen naar aanleiding van dit onderzoek dat het vroege gebruik van ^{18}FDG PET in het traject van stadiumbepaling (direct na het maken van de longfoto) de kwaliteit van stadiumbepaling niet beïnvloedt en dit proces niet vereenvoudigd vergeleken met de traditionele manier.

Referentie

1. Dijck JAAM van, Coebergh JWW, Siesling S Trends of cancer in the Netherlands 1989-1998. Utrecht: Vereniging van Integrale Kankercentra; 2002.

Dankwoord

Dankwoord

Uiteindelijk ligt het proefschrift dan toch voor u. Het verboden P-woord mag weer uitgesproken worden. Dit proefschrift is tot stand gekomen dankzij de inspanningen en samenwerking van heel veel collega's. Het was een voorrecht om op twee afdelingen (nucleaire geneeskunde en longziekten) te mogen werken en met veel collega's uit andere ziekenhuizen te kunnen samenwerken. Een belangrijk deel van dit proefschrift is klinisch gericht onderzoek, daar hebben veel patiënten belangeloos aan meegewerkt. Al deze mensen ben ik zeer dankbaar.

Professor dr. PE Postmus, beste Piet, erg blij was ik dat je me een kans gaf voor de opleiding tot longarts en de mogelijkheid om te promoveren. Op cruciale momenten tijdens het onderzoek maar ook binnen de opleiding tot longarts, gaf je me een zet in goede richting.

Professor dr. GJ Teule, beste Jaap, hoe druk je het ook had, er was eigenlijk altijd ruimte voor een filosofisch getint gesprek ergens tussendoor op weg van A naar B. Niet alleen voor de activiteiten op de werkvloer maar vooral ook voor bezigheden buiten het werk om heb je een brede belangstelling (zelfs de componist Kagel kreeg een kans). Dank voor al je inzicht.

Professor dr. OS Hoekstra, beste Otto, hoeveel tijd heb jij wel niet in dit proefschrift gestoken. En hoe vaak heb ik niet gehoord: "We doen het helemaal anders". Na heel wat re-analyses, herschrijven met behulp van jouw creativiteit ligt er nu toch maar een proefschrift! Je inzicht en bevlogenheid is erg aanstekelijk. Ik ben nog steeds blij met onze samenwerking.

Professor dr. EF Smit, beste Egbert, het kritisch denken is erin gestampt door jou. Bij elke vraag een wedervraag. Je scherpe maar ook praktische aanpak bij knelpunten was en is erg verhelderend. Dat geldt voor zowel de wetenschap als de kliniek. Daarnaast is het ook gewoon prettig om met jou samen te werken.

De leescommissie Prof. dr. H Groen, Prof. dr. M Boers, Prof. dr. J Vansteenkiste en dr. P Jager. U wil ik danken voor uw aandacht en de tijd die u genomen heeft voor dit proefschrift.

Alle collega's uit de perifere ziekenhuizen en het AZG zou ik willen bedanken voor hun bijdrage aan onder andere de POORT-studie. Hoe lastig het ook soms was, uiteindelijk hebben jullie 465 patiënten gerekruteerd. Dankzij jullie inspanning hebben we deze studie uit kunnen voeren.

De medewerkers van het IKA zou ik willen danken voor hun inzet voor de POORT-studie. Met name mevr G Smit en mw S de Ruiter, Gertruud en Stella, dank voor jullie grenzeloze inzet en jullie aangename gezelschap tijdens de eindeloze controles van de database. Jullie hebben mij flink bijgespijkerd ten aanzien van databases.

De eerste mijlpaal, mijn eerste internationale publicatie samen met drs. P van Velthoven en drs. J van den Bergh. Dank voor jullie hulp.

Mw EW Grijseels, mw C Uyl de Groot en dhr P Verboom, beste Els Carin en Paul. Dank voor jullie vakkundige kosten analyses.

Computer tomografie beelden beoordelen, dat doe je met drs R Golding, in de weekeinden nog een keer alle beelden bekijken. Erg leerzaam. Bedankt.

Dr. E Comans, Emile, fijn dat je altijd bereid was en bent een PET scan te beoordelen.

Drs J van Mourik en Dr MA Paul wil ik danken voor hun adviezen en vertrouwen ten aanzien van de POORT studie.

Dhr H van Tinteren, beste Harm, wat fijn dat je me eigenlijk altijd gerust kon stellen als ik dacht dat ik weer iets statistisch relevants over het hoofd gezien had. De tijdstippen dat ik je met mijn bedenksels lastig gevallen heb, waren volgens mij statistisch significant vaker net voor jou vakantie. Ondanks dat toch altijd een pasklaar antwoord.

Dr. A van Lingen, beste Arthur, computers en dergelijke blijven grote mysteries voor mij, blij dat jij met regelmaat een mysterie hebt op kunnen lossen. Jouw relativeringsvermogen en jouw steun in de rug zijn altijd erg welkom.

Veel administratief werk is verricht door mw A Kalwij, mw C Karga (PET centrum) en mw A Kijk in de Vegte. Beste Amanda, Cemile en Anny dank voor al jullie hulp bij al die klusjes die toch maar gedaan moeten worden.

De andere opleiders in het VUMC, dr. G Sutedja, dr. A Boonstra, dr. A Vonk Noordegraaf en dr. P Kunst wil ik danken voor hun geduld. Opleiden lijkt mij niet makkelijk... al die eigenwijze assistenten. Over enkele maanden kom ik terug voor de finishing touch!

Mijn kamergenoten Ben Venmans, Ton van Boxem, Cas Colder, Nel Hagen, Roderick Breuer, Igor Krivokuca, Hugo Rutten, Michel van de Heuvel, Wouter Jacobs, Hes Brokx, Arjen Winkel en Arifa Moons wil ik danken voor hun geduld wanneer ik weer stuiterend de kamer binnen kwam. Lekker klagen maar vooral ook veel gelachen.

De medewerkers van de behandelkamer en longfunctie, Linda Penninx, Hanneke Motman, Andy, Herman Groepenhoff, Gwenda Commandeur, Joost Admiraal en Wilma de Ruiter. Ik ben blij dat jullie zo nu en dan een onvolledig ingevuld formulier door de vingers zagen. Door

drukte moet je soms aan triage doen. Inmiddels weer op volle sterkte, zodat ik straks alles goed kan maken.

Het PET centrum met mw S van Baalen, dhr R Coopmans en destijds dhr B Hoving. Beste Suzette, Rob en Bas, dank voor jullie harde werken en de altijd lieve aandacht voor de patiënten.

Zeven c (onder leiding van Marieke Catz en Natasja Kok) en acht b (onder leiding van Frits Stricker) mag ik niet vergeten. De longafdelingen, met een geweldig team van verpleegkundigen. Het is een plezier om samen met jullie de afdeling te "runnen".

Collegae onderzoekers Bart Wittgen, Bas Holverda, Roald Roeleveld, Tji Tan, Jan-Willem Lankhaar, Remco van den Berg en Serge van Wolferen. De vrolijke noten van de afdeling. Onderzoek is taai en daarom heb je elkaar nodig om veel en hard te lachen, elke keer weer.

De collega's van mijn huidige werkplek het OLVG afdeling longziekten wil ik bedanken voor de vrije tijd die ik kreeg t.a.v. de afronding van dit proefschrift en voorbereiding op de promotie.

Graag wil ik huidige opleiders ten aanzien van het perifere deel van de opleiding tot longarts in het OLVG, dr. B Kwa, dr. W Hamersma, dr. F Krouwels, drs. F Toben, drs. K Liesker danken voor hun hartelijke en open ontvangst als nieuwe assistent.

De opleiders en alle collega assistenten van de afdeling interne geneeskunde uit het Sint Antonius ziekenhuis te Nieuwegein wil ik danken voor de twee erg leuke en leerzame jaren. Ik hoop dat dit proefschrift de vraag van dr. H Haanen ("wat moet je nou met zo'n PET scan?") beantwoord.

Een van de weinige mensen die ik nooit gestresst heb gezien is Chris Bor, hij heeft de gehele lay-out verricht. Dat gaf erg veel rust. Bedankt.

Een van mijn lieve vrienden Alex de Jong heeft een prachtig schilderij gemaakt. Een detail hiervan heb ik gebruikt voor de omslag. Met het woord wetenschap ben je aan de slag gegaan, met resultaat. Dank.

Mijn paranimf dr. C Hoekstra, Corneline, is altijd trouw geweest. Ondanks thuisdrukte genoeg tijd om even naar Utrecht te komen. Inmiddels geen directe collega meer, maar gelukkig nog steeds een heel goede vriendin.

Familie, vriendin en paranimf, drs. FE Jansen. Floortje, ik ben blij dat je naast alle andere life-events ook dit moment van dichtbij mee wil maken. Binnenkort sta ik jou terzijde.

Al mijn lieve vrienden, vriendinnen en familie, voor veel van jullie was ik de afgelopen periode niet erg bereikbaar. Het goede voornemen is er, om daar verandering in te brengen.

Jaap Tuijn, veel geoefend, herhaald, verbeterd, gelachen om mijn uitspraak. Samen de stress momenten doorstaan. Altijd bereid te helpen. Fijn dat je naast me stond.

Robert en Esther Herder, met een relativeringsvermogen als geen ander. Heel prettig wanneer ik door de bomen het bos niet meer zag.

Mijn lieve ouders, jullie krijgen alle dank voor alles wat jullie mij gegeven hebben. De mogelijkheid om te kunnen studeren, het tomeloze vertrouwen en alle andere wijsheden. Zonder jullie zou ik hier niet staan.

List of abbreviations

List of abbreviations

AUC	area under the curve
AIC	Akaike's information criterion
BGO	bismuth germanium oxide
CI	confidence interval
CT	computed tomography
DU	diagnostic understanding
ECOG	eastern co-operative oncology group
^{18}FDG	^{18}F-fluorodeoxyglucose
FU	follow-up
IQR	inter quartile range
MCA	medical centre Alkmaar
MRI	magnetic resonance imaging
NCR	national cancer registry
NPV	negative predictive value
NSCLC	non-small cell lung cancer
OSEM	ordered subset expectation maximisation
PALGA	pathological anatomical national register
PET	positron emission tomography
PPV	positive predictive value
ROC	receiver operating characteristic
SCLC	small cell lung cancer
SD	standard deviation
SPN	solitary pulmonary nodule
SUV	standardised uptake value
TC	treatment choice
TNM	tumour node metastasis
T/N ratio	tumour normal tissue ratio
TWU	traditional work-up
VUMC	VU university medical centre